RODOLPHE GHIGLIONE
BENJAMIN MATALON

Les enquêtes sociologiques

Théories et pratique

ARMAND COLIN

723 GHI

LANCHESTER LIBRARY
3 8001 00193 2148

LES ENQUÊTES SOCIOLOGIQUES

Théories et pratique

Collection U, série «Sociologie» dirigée par Henri Mendras

Philippe BESNARD, *Protestantisme et capitalisme* (U2)
Pierre BIRNBAUM, François CHAZEL, *Sociologie politique — Textes* (U2)
Pierre DESMAREZ, *La sociologie industrielle aux États-Unis*
W. DOISE, J.-C. DESCHAMPS, G. MUGNY, *Psychologie sociale expérimentale*
Claude DUBAR, *La socialisation : construction des identités sociales et professionnelles*
François DUPUY, J.-C. THŒNIG, *Sociologie de l'administration française*
Olivier GALLAND, *Sociologie de la jeunesse*
Rodolphe GHIGLIONE, Benjamin MATALON, *Les enquêtes sociologiques*
R. GHIGLIONE, J.-L. BEAUVOIS, C. CHABROL, A. TROGNON, *Manuel d'analyse de contenu*
Maurice HALBWACHS, *Morphologie sociale* (U2)
Bernard KAYSER, *La renaissance rurale*
Basile KERBLAY, *La société soviétique contemporaine*
Judilh LAZAR, *Sociologie de la communication de masse*
Yannick LEMEL, *Stratification et mobilité sociale*
Patrice MANN, *L'action collective, Mobilisation et organisation des minorités actives*
Henri MENDRAS, *Éléments de sociologie*
Éléments de sociologie — Textes (U2)
Henri MENDRAS, Michel VERRET, *Les champs de la sociologie*
Henri MENDRAS, Michel FORSÉ, *Le changement social*
Jean PENEFF, *La méthode biographique*
Jean-Pierre POITOU, *La dissonance cognitive*
Jean-Daniel REYNAUD, *Les règles du jeu. L'action collective et la régulation sociale*
Martine SÉGALEN, *Sociologie de la famille*
Denis SEGRESTIN, *Sociologie de l'entreprise*
Georges TAPINOS, *Éléments de démographie*
Michel VERRET, *Le travail ouvrier*

Rodolphe Ghiglione
Benjamin Matalon

LES ENQUÊTES SOCIOLOGIQUES
Théories et pratique

Quatrième édition

ARMAND COLIN

Ce logo a pour objet d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, tout particulièrement dans le domaine universitaire, le développement massif du « photocopillage ».
Cette pratique qui s'est généralisée, notamment dans les établissements d'enseignement, provoque une baisse brutale des achats de livres, au point que la possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.
Nous rappelons donc que la reproduction et la vente sans autorisation, ainsi que le recel, sont passibles de poursuites. Les demandes d'autorisation de photocopier doivent être adressées à l'éditeur ou au Centre français d'exploitation du droit de copie : 3, rue Hautefeuille, 75006 Paris. Tél. : 43 26 95 35.

Tous droits de traduction, d'adaptation et de reproduction par tous procédés réservés pour tous pays.

Toute reproduction ou représentation intégrale ou partielle, par quelque procédé que ce soit, des pages publiées dans le présent ouvrage, faite sans l'autorisation de l'éditeur est illicite et constitue une contrefaçon. Seules sont autorisées, d'une part, les reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective et, d'autre part, les courtes citations justifiées par le caractère scientifique ou d'information de l'œuvre dans laquelle elles sont incorporées (art. L. 122-4, L. 122-5 et L. 335-2 du Code de la propriété intellectuelle).

Lo 5978 Coventry University

4e édition, 4e tirage

© Armand Colin Éditeur, Paris, 1977, 1985
ISBN : 2-200-21635-1

Armand Colin Éditeur - 5, rue Laromiguière - 75241 Paris Cedex 05

Chapitre 1

PROBLÈMES GÉNÉRAUX

1.1. DÉFINITIONS ET GÉNÉRALITÉS

L'enquête est un des instruments les plus largement utilisés par les psychologues sociaux et les sociologues. Des études de marché jusqu'aux recherches purement théoriques en passant par les sondages d'opinion, il est peu d'investigations psychosociologiques ou sociologiques empiriques qui ne s'appuyent, partiellement ou en totalité, sur des informations recueillies par enquêtes. Pourtant, si on dispose de quelques ouvrages techniques sur la manière de réaliser une enquête, ouvrages d'ailleurs étonnamment peu nombreux eu égard à l'importance de l'instrument, il n'en existe pratiquement pas qui traite dans leur ensemble les problèmes à la fois théoriques et pratiques que l'usage de ces techniques soulève [1].

Notre ambition est de passer en revue les différents aspects techniques des enquêtes, et de fournir ainsi au lecteur praticien des indications utiles. Mais ceci ne signifie pas que l'on se consacrera uniquement à l'étude d'une pratique, ce qui est fréquent dans ce

1. Il n'y a guère que Lazarsfeld et ses collaborateurs qui, depuis de nombreuses années, réfléchissent au développement des techniques d'enquête et à leurs aspects théoriques. En France, les travaux de R. Boudon et de son équipe (le Groupe d'Etudes des Méthodes de l'Analyse Sociologique) sont dans le même cas.

type d'ouvrage ; on abordera également les problèmes théoriques ou épistémologiques posés par la méthode.

Il est difficile de parler de « l'enquête » en général, non seulement parce qu'il en existe différents types, mais surtout parce que sa pratique exige le recours à différentes techniques, qui soulèvent chacune des questions spécifiques : méthodes de sondage, entretiens libres, échelles d'attitude, analyse de contenu, analyse statistique, etc. Il n'y a donc pas de « théorie des enquêtes », mais un ensemble de questions théoriques, épistémologiques, méthodologiques très diverses.

Réaliser une enquête, c'est *interroger* un certain nombre *d'individus* en vue d'une *généralisation*. Cette définition, relativement arbitraire et certainement contestable comme toute définition, vise à distinguer l'enquête des autres méthodes sociologiques. D'abord par le fait qu'il s'agit d'interrogation, elle se distingue à la fois de l'observation, où l'intervention du chercheur essaie d'être minimale, et de l'expérimentation, où le chercheur, au contraire, crée et contrôle la situation dont il a besoin. Ensuite, prenant comme unité d'observation, et donc d'analyse, des individus, l'enquête se distingue des différentes méthodes sociologiques qui portent directement sur des unités plus vastes, groupes, classes ou institutions. Enfin, il s'agit d'interrogation en vue d'une généralisation : ce ne sont pas les individus dans ce qu'ils ont de personnel qui nous intéressent, comme c'est le cas dans un entretien-diagnostic ou un entretien d'embauche, mais la possibilité de tirer de ce qu'ils disent des conclusions plus larges.

Une enquête consiste donc à susciter un ensemble de discours individuels, à les interpréter et à les généraliser. Les problèmes théoriques et méthodologiques soulevés par sa pratique et son usage peuvent être rattachés à ces caractéristiques et ne peuvent être analysés que par rapport à elles.

Les discours qui constituent la « matière première » de l'enquête ne sont pas spontanés ; ils ne sont pas produits dans un vide social qui en garantirait l'objectivité. Ils ont été obtenus dans une situation très particulière d'interaction sociale, situation structurée en grande partie, mais pas uniquement, par la relation qui s'établit entre l'enquêteur et l'enquêté. Comme, d'ailleurs, dans toutes les situations provoquées, les situations expérimentales par exemple, il faut tenir compte du fait qu'on n'a aucune raison d'admettre que le sujet se pliera passivement aux consignes et qu'il délivrera directement « la vérité », ni même, plus modestement, « sa » vérité. Consciemment ou non, il ne nous dit que ce qu'il *peut* et *veut* nous dire. Ceci est déterminé par la *représentation* qu'il se fait de la situation, par ses propres *objectifs*, qui ne coïncident pas nécessairement avec ceux du chercheur. Un tel discours ne peut être interprété que par référence aux conditions dans lesquelles il a été produit. Cela est vrai de tout discours, mais il faut renoncer à l'illusion qui consisterait à croire que, grâce par exemple à la volonté de neutralité de l'enquêteur ou à la standardisation des questions, la situation d'enquête n'exercerait aucune influence sur les réponses des sujets.

Tout le matériel qu'on recueille dans cette situation, et sur lequel on travaillera, est exclusivement verbal, ce qui pose le problème du *sens*. Ce problème se manifeste à travers les différentes formes d'incompréhension qui peuvent se produire entre les différentes catégories d'enquêtés, les enquêteurs, et le chercheur qui interprète les discours obtenus.

Au problème du sens, on peut rattacher tout ce qui a trait à la formulation des questions et à l'analyse de contenu des réponses ; ceci implique de tenir compte des différences d'usage du langage dans les différentes catégories sociales.

C'est également dans ce cadre qu'on peut se demander quelles raisons on a de considérer une réponse ou un ensemble de réponses comme un *indicateur* d'une variable de portée plus générale que les réponses effectivement obtenues (comme c'est le cas, par exemple, lorsqu'on infère une attitude à partir de l'énoncé d'un ensemble d'opinions particulières).

Au moment de l'interprétation, nous disposons d'une *collection de discours individuels*, à partir desquels il nous faut tenir un discours *unique* : rapport de fin d'étude, publication dans une revue scientifique, ou toute autre forme d'analyse des résultats. Cela pose le double problème de *l'agrégation* des réponses individuelles et de leur *généralisation*.

Le mécanisme de *l'agrégation* amène à se demander quels sont les groupes d'individus dont on peut confondre les réponses, à propos desquels il est légitime de tenir un discours unique ; il renvoie donc à des problèmes de typologie (du point de vue formel, les classes sociales, par exemple, peuvent être considérées comme constituant une typologie). D'autre part, on rencontre les problèmes d'inférence statistique, qui précisent la généralité possible du discours résultant, d'où l'attention portée aux méthodes d'échantillonnage. C'est aussi dans ce cadre théorique qu'on peut s'interroger sur la pertinence d'une sommation simple de réponses individuelles pour la compréhension des phénomènes sociaux, autrement dit, pour parler un peu rapidement, de la spécificité des niveaux psychologique et sociologique.

L'interprétation des résultats, comme dans toute recherche méthodique, mais de façon peut-être plus difficile parce que l'arrière-plan théorique est souvent plus faible ici qu'ailleurs, soulève le problème de la *causalité*.

Ce problème, ou plutôt celui des inférences causales, apparaît aussi bien dans l'emploi des méthodes informelles utilisées le plus souvent pour l'interprétation des entretiens libres que lorsqu'on essaie de tirer des conclusions d'une analyse statistique portant sur des concomitances qui s'expriment par des différences ou des corrélations.

Les sujets interrogés, surtout lorsqu'on les laisse s'exprimer librement comme dans des entretiens non directifs, n'ont à leur disposition qu'une gamme restreinte de concepts et un vocabulaire limité (en tout état de cause, ce ne sont le plus souvent ni les concepts, ni le vocabulaire du sociologue ou du psychologue). Ils se heurtent donc à des limites objectives qui n'ont peut-être aucune relation avec le phénomène lui-même, mais avec les représentations que les sujets en ont, et surtout celles qu'ils peuvent communiquer facilement, donc fortement déterminées par ces instruments intellectuels.

Le problème se pose différemment lorsque c'est le chercheur lui-même qui impose ses concepts et son vocabulaire, comme dans des questionnaires fermés. Ce faisant, il risque de placer le sujet devant sa propre conception du phénomène à étudier et l'obliger à accepter son propre découpage. Le sujet va peut-être l'accepter, mais il se peut, même dans ce cas, que ce ne soit pas le point de vue qu'il aurait adopté de lui-même.

Ces différents problèmes théoriques sont très généraux et semblent, à première vue, très éloignés des difficultés concrètes que rencontre quotidiennement le praticien. Une

réflexion sommaire nous montre toutefois que chacun d'entre eux renvoie à des problèmes pratiques plus ou moins clairement reconnus, et que cette démarche critique, théoriquement située, permet de mieux les poser.

Par exemple, reconnaître la spécificité de la *situation d'enquête* en tant qu'*interaction sociale*, et son influence sur la production du discours, renvoie aux différentes manières de définir, on pourrait dire de créer, cette situation par le choix des enquêteurs, la manière dont ceux-ci se présentent aux futurs sujets et leur présentent l'objet de l'enquête, l'image d'eux-mêmes et de leurs attentes qu'ils communiquent, volontairement ou non, aux enquêtés par la consigne, par leur attitude générale, leurs réactions même involontaires, etc. C'est par conséquent poser aussi la question de tous les biais qui peuvent affecter les réponses, biais liés à des normes, à des stéréotypes, à des représentations diverses, et auxquels il faut penser au moment de l'interprétation.

Pour terminer ce survol des problèmes généraux, mentionnons également une objection qui est souvent faite surtout par les sociologues à l'usage des enquêtes. On reproche à la méthode de ne saisir que des individus artificiellement isolés par le mode même d'interrogation, et de réduire ainsi inévitablement les problèmes à leur aspect individuel, alors que des analyses plus complètes, et non biaisées, devraient porter sur des entités telles que les groupes, les classes sociales, les organisations, les institutions, etc. L'enquête, en particulier l'enquête par sondage auprès d'un échantillon considéré comme représentatif, qui procède par sommation de réponses individuelles, amènerait à une « psychologisation » des problèmes. Les processus sociaux réels se déroulent non entre des individus, monades isolées, mais par des interactions entre des groupes ; on ne saisirait, par des enquêtes individuelles, qu'un aspect très partiel, et artificiellement découpé, de ce processus. Certes une enquête peut toujours apporter des résultats, mais on n'y retrouverait que ce qu'on y avait mis au départ : des réponses individuelles, et donc un certain niveau d'analyse, à l'exclusion des autres.

Cette objection n'est pas sans fondement, et il est clair qu'il y a des problèmes qui ne relèvent pas de l'enquête. Que la méthode soit souvent appliquée, même dans de tels cas, c'est vrai aussi, comme il est vrai que ce choix, à première vue purement technique, peut, en définitive, se révéler idéologique : sous le simple aspect d'un choix de méthode, on peut être amené à privilégier un niveau d'analyse individuel, omettant ainsi de reconnaître que les « acteurs sociaux » peuvent être autres que des individus. L'importance de cet argument exige qu'on en apprécie exactement la portée, afin de mieux cerner le champ où le recours à l'enquête reste justifié.

Pour fixer les idées, prenons comme premier exemple les sondages d'opinion, largement pratiqués et diffusés. Laissons pour le moment de côté le problème de savoir ce qu'est cette « opinion », s'il s'agit véritablement d'un objet scientifique, et admettons provisoirement que chacune des questions posées a vraiment un sens, et le même, pour tous les sujets interrogés (nous reviendrons plus loin sur ce point, évidemment capital). Bornons-nous à nous demander ce que signifie la sommation des réponses aux mêmes questions de 2 000 personnes choisies de manière que les proportions d'hommes et de femmes, de cadres et d'ouvriers, de citadins et de ruraux, de riches et de pauvres, de jeunes et de vieux, etc., soient celles de l'ensemble des Français. Cette manière de

procéder est effectivement très pauvre. L'« opinion publique » qui a une signification politique ou sociologique n'est pas simplement une juxtaposition de réponses individuelles, mais représente l'aboutissement d'un phénomène dynamique complexe, où interviennent des processus d'influence et de pouvoir, où tous les individus n'ont pas le même poids, où les leaders et les moyens de communication de masse jouent un rôle essentiel, la position exprimée par ces derniers étant à son tour déterminée, entre autres, par leur situation dans le système social. Tout cela, le sondage d'opinion le néglige, pour ne chercher à donner, comme le disent prudemment les spécialistes, qu'une « photographie de l'opinion à un moment donné ». Certes, il fournit souvent des indications intéressantes, surtout lorsqu'on peut comparer dans le temps des sondages successifs, mais il ne permet guère d'analyser l'opinion comme étant à la fois produit et facteur d'une dynamique sociale, ni de prévoir des comportements [1].

Nous emprunterons notre second exemple à l'étude de P. Bourdieu sur la photographie, qui s'ouvre par une critique violente des études de motivation effectuées sur le sujet, auxquelles il reproche, précisément, leur « psychologisme ». Pour lui, montrer simplement que la quantité de pellicule consommée par an augmente avec le revenu, renvoie à un modèle explicatif simpliste, selon lequel il y aurait une « envie » universelle (ou du moins répartie de la même façon dans l'ensemble de la population) de photographier, envie qui serait simplement freinée par des revenus insuffisants. Pour Bourdieu, et même si l'on analyse « en profondeur » cette envie de photographier, c'est se condamner à ne rien comprendre à la pratique photographique, qui varie considérablement, et pas seulement en quantité, d'un groupe social à l'autre. Selon le groupe auquel on appartient, on ne photographiera pas les mêmes choses, et pas dans les mêmes occasions ; les critères de la « bonne photo » varieront, ainsi que ce qu'on fait de la photo une fois prise et développée. Les différences entre catégories sociales sont telles que chercher à les saisir par un questionnaire unique serait complètement illusoire et réduirait le phénomène étudié à ce qui est commun à tous les groupes, c'est-à-dire peu de choses, et pas le plus important. Ce qu'il faut faire, et c'est effectivement ce que Bourdieu et son équipe ont réalisé, c'est étudier séparément les différents groupes, chacun avec la méthode la mieux adaptée : analyse des albums de photos des habitants d'un village, observation participante dans le club de photographes amateurs d'une grande industrie, entretiens auprès de professionnels, etc. On obtient ainsi une série de monographies qui, juxtaposées, permettent de saisir ce qu'est, sociologiquement, la photographie, sans chercher à aller au-delà de ce rapprochement pour réaliser une synthèse qui n'aurait guère de sens.

Ces objections paraissent parfaitement valables, et le danger qu'elles soulignent est réel. Toutefois, en tenir compte n'amène pas nécessairement à renoncer aux enquêtes, mais plutôt à les mettre à leur juste place, qui est d'obtenir des réponses d'individus, réponses dont le plus souvent on ne peut pas se passer. Après tout, ce sont, en définitive, les individus, et non les groupes, qui agissent, qui ressentent, qui préfèrent. Que

---

1. Sauf parfois des votes ; mais c'est un cas privilégié où il y a effectivement correspondance entre les conditions d'une consultation électorale et la situation de sondage (choix dont les conditions sont fixées d'avance, décisions isolées, poids égal de tous les individus, etc.).

ces actions, ces affects, ces préférences soient socialement déterminés, que les actions importantes ne soient pas le fait d'individus isolés, c'est évident ; il n'empêche que ces phénomènes ne peuvent être saisis qu'à ce niveau. C'est au moment de la synthèse, au moment où l'on met ensemble toutes ces données individuelles pour en tirer des conclusions et des généralisations qu'on risque de commettre des erreurs graves. Si l'on admet sans réflexion qu'il est légitime de sommer sans plus des réponses, qu'aucune autre information n'est nécessaire, alors effectivement on aura émietté la dynamique sociale, et ce qui en restera sera sans intérêt. Mais ce n'est pas la seule voie possible. Et même si l'on pense avoir de bonnes raisons de procéder ainsi, il faut se souvenir que les phénomènes sociaux ne se réduisent que rarement à ce qu'on peut tirer d'une enquête seule : d'autres approches, complémentaires, sont souvent indispensables. Dans les deux exemples que nous avons cités, à un certain moment, des enquêtes se sont révélées nécessaires. L'important est de bien préciser leur objet.

Revenons à l'exemple des sondages d'opinion. Il est clair que cela n'a pas de sens d'« interroger » des groupes sociaux, même si nous pensons que c'est à ce niveau que se situent les acteurs réels des processus qui nous intéressent. Se contenter des leaders et des moyens de communication de masse, c'est oublier que les individus ne les suivent pas passivement et se condamner à ne pas pouvoir analyser leur influence réelle et la manière dont cette influence s'exerce. De plus, sans interrogation directe des intéressés, nous ne pouvons pas savoir quelles sont leurs représentations, leurs normes, leurs cadres de référence, qui sont certes des produits sociaux, mais qu'il faut bien, à un moment ou à un autre, saisir en eux-mêmes.

De ces réflexions, deux idées nous semblent particulièrement importantes à retenir. D'abord, les phénomènes sociaux ne peuvent pas être réduits à ce qu'on peut tirer de ce qu'en disent des individus : d'autres informations sont nécessaires et peuvent être recueillies par les différentes méthodes que sont l'observation, l'expérimentation, l'analyse de traces ou de documents divers. Ensuite, il ne faut pas confondre l'enquête, qui est une méthode très générale, avec une de ses modalités, la plus courante il est vrai : le sondage auprès d'un échantillon représentatif. Utiliser des personnes bien placées comme des informateurs, interroger d'autres personnes de qualité relève d'une autre démarche, mais qui est souvent indispensable pour reconstituer et analyser un processus social.

1.2. ENQUÊTE, OBSERVATION, EXPÉRIMENTATION

Si l'on veut bien admettre que tout chercheur analyse une situation à l'aide d'une ou de plusieurs des quatre méthodes que constituent :

— l'observation,
— l'enquête,
— l'expérimentation,
— l'étude des « traces »,

et que le choix, au moins dans les sciences sociales, n'est pas plus étendu, trois ques-

tions se posent. La première est celle de la définition de chacune de ces méthodes, la seconde celle de leur interaction, la troisième celle des conditions du recours à l'enquête.

1.2.1. Définitions

L'observation peut être définie comme un regard porté sur une situation sans que celle-ci soit modifiée. Regard dont l'intentionnalité est de nature très générale et agit au niveau du choix de la situation et non au niveau de ce qui *doit* être observé dans la situation, le but étant le recueil des données afférentes à la situation.

L'enquête, quant à elle, peut être définie comme une interrogation particulière portée sur une situation comprenant des individus, et ce, dans un but de généralisation. Dans ce cas, le chercheur intervient en posant des questions, mais sans avoir le désir explicite de modifier la situation dans laquelle il agit en tant que tel (Une telle définition ne doit pas faire penser que les auteurs ont la naïveté de croire que les interrogations du chercheur, ni même la présence de l'observateur, ne modifient pas la situation. Ils veulent tout simplement marquer que telle n'est pas l'intention du chercheur. Ils se réservent de revenir sur ce point et de l'expliciter par la suite.

L'expérimentation peut être définie comme une interrogation particulière portée sur une situation créée et contrôlée par le chercheur (là encore si la notion de création d'une situation n'est pas ambiguë, il n'en va pas de même de la notion de contrôle, nous y reviendrons. Cette interrogation est à but de vérification (d'une hypothèse ou d'un corps axiomatique dont l'hypothèse est le produit).

L'étude des traces peut être considérée comme une forme d'observation différée, qui par nécessité ne saisit pas directement le phénomène intéressant, mais uniquement certaines de ses conséquences. On peut regrouper dans cette catégorie aussi bien l'analyse de documents, de statistiques officielles, que de véritables traces matérielles. C'est le type des méthodes dites non réactives : puisque le chercheur n'intervient qu'après que le phénomène se soit produit, il ne peut évidemment pas le perturber. Mais les processus de production et de conservation des traces peuvent être complexes, et leur lien avec le phénomène étudié peut donc être faussé. De plus, certaines traces, comme par exemple les statistiques officielles, sont elles-mêmes des produits sociaux, dont la validité peut être questionnée.

1.2.2. Relations entre les méthodes

Pour préciser ce que nous entendons par « enquête », nous avons été obligés de situer celle-ci par rapport aux autres grandes classes de méthodes de la psychologie sociale, et de décrire les types idéaux plus que les démarches effectives de ceux qui les pratiquent, ce qui nous a conduits à mettre plus particulièrement l'accent sur ce qui les distingue les unes des autres. Néanmoins, les points qu'elles ont en commun sont nombreux et divers. Les sous-estimer, et par-là fixer leurs oppositions, aurait pour conséquence d'accentuer les clivages à l'intérieur de la psychologie sociale, clivages qui, pour être actuellement bien réels, ne nous semblent pas épistémologiquement fondés. C'est pourquoi il

nous paraît opportun de dire quelques mots sur ce que ces méthodes ont en commun. Quant aux différences, qui sont d'ailleurs peut-être plus visibles, nous aurons l'occasion d'y revenir en détail lorsque nous parlerons du choix entre différentes méthodes, des apports et des limites spécifiques de chacune.

Ces points communs se situent à des niveaux très différents les uns des autres. En premier lieu, il ne faut pas oublier que l'utilisation de ces méthodes relève, en principe, d'une même visée scientifique, d'un même souci d'*administration de la preuve* ; il y a donc, de ce fait, une logique sous-jacente qui leur est commune. En particulier, ce que nous appellerons l'*attitude expérimentale*, qu'il faut distinguer de l'expérimentation en laboratoire qui en est une application particulière, se retrouve dans la démarche du chercheur qui conçoit et interprète une enquête : le souci d'expliciter les hypothèses que l'on cherche à vérifier, la réflexion critique sur les conditions de leur vérification et, pour y arriver, le souci de contrôler les facteurs, de ne comparer que ce qui est comparable. Mais, alors que l'expérimentateur cherche à *créer* les conditions qui l'intéressent, ce qui lui donne, en principe, les possibilités de les faire varier systématiquement en fonction de ses hypothèses, celui qui procède par enquête *choisira* son terrain ou ses populations pour pouvoir effectuer les analyses dont il a besoin pour atteindre ses objectifs. Et lorsqu'il procède au traitement statistique des résultats qu'il a recueillis, les procédures qu'il utilise relèvent, elles aussi, du même mode de pensée que la constitution des situations expérimentales et leur analyse [1]. De même, l'observation n'est jamais réductible à un simple enregistrement : là aussi, le chercheur choisit ses unités, les traits pertinents qu'il va relever, etc., et ces choix, en dernière analyse, relèvent eux aussi de la même logique que la mise au point d'une expérience de laboratoire.

La similitude entre enquête et expérimentation peut être poussée encore plus loin : au lieu de se contenter d'interroger un échantillon représentatif, au sens statistique, de la population étudiée, on peut constituer des groupes, les sous-échantillons, conçus en fonction d'un plan expérimental, comme on le fait en laboratoire (cf. 2.1). Pour étudier, par exemple, l'influence simultanée de la catégorie sociale et de l'appartenance politique sur un certain comportement, on pourra constituer des sous-échantillons égaux d'individus appartenant à chaque catégorie sociale et à chaque groupe d'appartenance politique, selon un plan factoriel, chaque catégorie sociale étant associée aussi souvent que les autres à chaque appartenance politique, ce qui assure, pour l'ensemble des personnes étudiées, l'indépendance entre les deux facteurs, et par conséquent la possibilité d'évaluer séparément l'influence de chacun, et celle de leur interaction. La logique des plans d'expérience se retrouve, identique à ce qu'elle est au laboratoire.

Pour certains, c'est ce type de démarche qui, seul, peut avoir un sens psychologique ou sociologique, puisque, au lieu de créer des situations qui ont toujours un aspect artificiel, on peut étudier l'influence de caractéristiques essentielles, constitutives des individus et non extérieures à eux, comme l'âge, le sexe ou l'appartenance à une catégorie sociale.

Ces méthodes, toutefois, malgré tout leur intérêt, ne reproduisent pas la totalité des

---

1. Nous reviendrons plus loin sur le fait que des étapes logiquement équivalentes se situent à des moments différents selon qu'il s'agit d'expériences ou d'enquêtes.

caractéristiques de la méthode expérimentale. En effet, au moins depuis R.A. Fisher, on sait que, pour s'assurer que les différences observées d'un groupe à l'autre puissent être sans ambiguïté attribuées au facteur qui distingue ces groupes, il est indispensable que les sujets aient été affectés de façon strictement aléatoire à chaque condition expérimentale, à chaque groupe du plan. Sinon, on pourra toujours objecter que l'effet observé est peut-être dû à une variable parasite corrélée avec une des variables expérimentales, et qu'on n'a pas réussi, ou pas pensé, à contrôler. C'est cette « randomisation » qui, avec l'orthogonalité des facteurs, c'est-à-dire leur indépendance, assure à la preuve expérimentale son caractère contraignant.

Or, cette randomisation n'est évidemment pas possible avec les groupes pris en considération dans les enquêtes : on ne peut pas attribuer aléatoirement aux sujets interrogés une certaine catégorie sociale, ni une appartenance politique. Les groupes constitués sur la base de telles caractéristiques, même s'ils sont conçus en fonction d'un plan expérimental irréprochable, ont toutes les chances de différer selon de nombreuses autres variables (âge, revenu, attitudes, personnalité, etc.), dont on sait qu'elles affectent de nombreux comportements. Les conclusions de ces observations seront d'autant plus incertaines que toutes les variables importantes, considérées comme « explicatives » du point de vue psychosociologique ou sociologique, présentent des intercorrélations complexes (cf. 7-2-2-2). D'autre part, le phénomène d'*auto-sélection,* qui constitue un processus particulier d'apparition de ces corrélations, vient rendre l'analyse encore plus difficile. Supposons que l'on veuille étudier l'influence de la télévision. Si on compare directement les comportements ou les attitudes de personnes qui la regardent ou non, les conclusions ont toutes les chances d'être biaisées par le fait que ce ne sont très vraisemblablement pas les mêmes personnes qui, spontanément, choisissent de regarder la télévision ou s'abstiennent.

A première vue, ces difficultés ne sont pas redhibitoires. On peut compliquer le plan expérimental en tenant compte explicitement des facteurs parasites, on peut constituer des échantillons aléatoires de personnes à qui on demandera de regarder, ou de ne pas regarder, la télévision durant une certaine période, indépendamment du fait qu'ils l'auraient ou non fait spontanément, ce qui nous éloignerait de l'enquête au sens strict. Toutefois, vu la complexité certaine des phénomènes sociaux, et l'absence de théorie nous indiquant quelles sont les variables pertinentes à prendre en considération [1], il est impossible d'admettre que des groupes choisis de façon non aléatoire seront bien équivalents.

Il résulte de tout cela que les conclusions, en termes de vérifications d'hypothèses portant sur des relations causales, restent toujours moins assurées dans les enquêtes que dans les expériences, où le caractère « extérieur » des situations créées en laboratoire assure leur indépendance par rapport à toutes les caractéristiques individuelles, et

1. Cette conséquence de l'absence de théorie constitue une des raisons du recours systématique aux techniques statistiques dans les sciences humaines : malgré tous les efforts qu'on peut déployer pour contrôler tous les facteurs, il reste toujours une variance résiduelle, non expliquée par eux, et qu'on ne peut qu'attribuer à des variables non prises en considération, malgré leur influence. Les physiciens n'ont pas besoin d'inférence statistique, sauf lorsqu'il s'agit de phénomènes que la théorie reconnaît comme intrinsèquement aléatoires. Ou alors, le nombre d'unités prises en considération est tel, comme en thermodynamique, que, malgré un modèle statistique, les observables constituent des variables certaines, sans autres fluctuations que les inévitables erreurs de mesure.

donc permet l'équivalence par randomisation. Il est toutefois possible, par des analyses fines faisant intervenir plusieurs facteurs simultanément, d'essayer d'éliminer suffisamment d'hypothèses concurrentes pour que celle qu'on cherche à vérifier en acquière une vraisemblance suffisante [1].

Un deuxième point commun est l'influence constante de la présence du chercheur, qu'il intervienne sous la forme d'un expérimentateur, d'un enquêteur ou d'un observateur, et même, à la limite, d'un ordinateur programmé pour présenter des stimuli et enregistrer les réponses. Dans chaque cas, le chercheur est présent dans la situation, il constitue un élément de la représentation que le sujet se fait de celle-ci. Certes, les degrés de cette intervention, et de la perturbation qu'elle peut entraîner, peuvent varier considérablement, et leurs conséquences peuvent être de gravité très diverse. Mais il faut prendre garde au fait qu'il ne suffit pas de n'avoir aucune intention de modifier une situation pour y parvenir effectivement.

Depuis quelques années, on a commencé à s'intéresser sérieusement à la psychosociologie de la situation expérimentale. De nombreuses recherches ont mis en évidence l'importance de la représentation que le sujet se fait de la situation dans laquelle il se trouve placé, de l'expérimentateur, de ses intentions et de ses attentes, ainsi que le rôle de la communication involontaire et inconsciente qui s'établit entre eux. Une analyse du même type peut être faite de la situation d'enquête et de la relation enquêteur-enquêté. Même l'observation - où pourtant, en principe, l'intervention du chercheur se veut nulle - n'échappe pas à ces difficultés. Effectivement on peut trouver des exemples où il est assez facile d'admettre que l'observateur n'influe pas du tout sur les comportements, en particulier lorsque l'on prend suffisamment de précautions pour que les sujets ne se sachent pas observés. Par exemple, un individu perdu dans une foule qu'il observe peut raisonnablement espérer ne pas la perturber par sa présence et les notes qu'il prend. Mais de tels cas restent peu nombreux. Et si on pense à des méthodes comme la recherche active *(action research)* ou l'observation participante, alors il faut bien reconnaître que l'intervention du chercheur devient un élément non négligeable, souvent même central, du processus qu'il veut étudier. Croire que le recours à l'enquête ou à l'observation permet de saisir des phénomènes plus « naturels » qu'en expérimentant est naïf. Le problème des généralisations possibles se pose dans chaque cas, bien qu'en des termes différents.

Les seules méthodes qui échappent à ces difficultés sont celles qui ne portent pas directement sur les individus eux-mêmes et sur leurs comportements, mais sur les « traces » de ceux-ci : écrits et documents « spontanés » divers, statistiques, comportements dérivés, etc. Toutefois, malgré leur intérêt et l'ingéniosité de certains de leurs utilisateurs, ces méthodes « non réactives » restent de portée limitée, et ne semblent pas encore en mesure de se substituer souvent aux méthodes classiques.

D'ailleurs, poursuivant plus loin le même type de réflexion, on a même remis en question la validité de documents habituellement considérés comme objectifs, puisque non suscités par la recherche elle-même, comme les statistiques officielles. Par exemple, on a montré, dans le cas des données statistiques sur le suicide, comme celles qu'a

1. Nous traitons en détail ce problème dans la dernière partie de ce livre.

utilisées Durkheim, que celles-ci sont elles-mêmes l'aboutissement de processus sociaux dans lesquels interviennent les représentations, les attitudes et les normes qui sont également des déterminants du suicide. Indépendamment de toute intention consciente de biaiser les statistiques, celles-ci ne constituent donc pas les données objectives qu'on y voit souvent. Elles sont tout autant à expliquer, et souvent par les mêmes facteurs, que les phénomènes qu'elles prétendent décrire.

Enfin, dans la réalité de la recherche ou des applications, les méthodes sont beaucoup moins distinctes qu'on pourrait le penser. On rencontre des intermédiaires difficiles à classer comme les expérimentations dites « en milieu naturel » ou en « vraie grandeur », ou des complémentarités comme le recours à des questionnaires ou à des entretiens post-expérimentaux. L'enquête peut souvent compléter utilement des observations qu'il est difficile d'interpréter sans avoir interrogé les intéressés. Certaines expériences de laboratoire portant sur les attitudes sont suivies de questionnaires de même nature que ceux utilisés au cours de certaines enquêtes, et qui visent à saisir la manière dont le sujet a perçu la situation.

1.2.3. Utilisation des enquêtes

Dans l'arsenal des méthodes dont dispose le psychologue social ou le sociologue, l'enquête occupe donc une position intermédiaire et, d'une certaine manière, ambiguë. Plus perturbante que l'observation, permettant moins bien que l'expérience en laboratoire de contrôler la situation et de susciter les conditions dont le chercheur a besoin, elle a en outre l'inconvénient très grave de s'appuyer exclusivement sur le langage, celui du chercheur et ceux des diverses catégories d'enquêtés, dont on n'a aucune raison d'admettre, a priori, qu'ils coïncident. Dans ces conditions, pourquoi faire des enquêtes, pourquoi est-ce, aujourd'hui, un des instruments privilégiés de la recherche en sciences sociales ?

La question semble naïve, et purement rhétorique. La pratique sociale ne constitue-t-elle pas la meilleure preuve de l'intérêt, de l'utilité de la méthode ? Certes, mais cela ne dispense ni d'une critique sérieuse de la technique, ni de la reconnaissance de ses limites, ni de l'explicitation des conditions dans lesquelles l'enquête apparaît effectivement comme le seul moyen d'obtenir certaines informations.

1. Le recours à l'interrogation est nécessaire chaque fois que l'on a besoin d'informations sur une grande variété de comportements d'un même sujet, comportements dont l'observation directe, même si elle était possible, demanderait trop de temps, ou serait complètement impossible parce qu'il s'agit du passé. Qu'on cherche à savoir dans quelles conditions s'est déroulée l'enfance du sujet, pour qui il a voté aux dernières élections, ou quand il a acheté des petits pois pour la dernière fois, il n'y a en général rien d'autre à faire que de le lui demander.

2. On recourt également à l'enquête lorsque l'observation directe de certains comportements supposerait une véritable intimité souvent pratiquement impossible et inacceptable d'un point de vue déontologique. Il ne reste alors qu'à se contenter des descriptions, des récits, que les sujets enquêtés voudront bien faire en réponse à nos questions.

3. Dans chacune de ces deux catégories de situations, l'enquête apparaît comme le substitut d'une observation trop difficile ou impossible. On est obligé de recourir également à l'interrogation pour saisir des phénomènes tels que les attitudes, les opinions, les préférences, les représentations, etc., qui ne sont pratiquement accessibles que par le langage, et qui ne s'expriment spontanément que rarement. Certes, on pourrait essayer d'en donner une définition strictement comportementale et, par exemple, chercher à inférer certaines préférences à partir de l'observation de choix réels, ou éventuellement des choix suscités en situation expérimentale. Mais, ce faisant, on s'interdirait de poser le problème des relations, et des désaccords éventuels, entre ces préférences et les comportements correspondants.

D'autre part, observer un comportement peut souvent ne pas être suffisant ; il faut également savoir comment le sujet l'explique, quelle *signification* il a pour lui, ce qu'on ne peut savoir que s'il nous l'explicite. Ceci ne veut d'ailleurs pas dire que nous nous en tiendrons au contenu manifeste de ce qu'il nous dit à ce propos. Cela peut amener, dans certains cas, à compléter des observations par des enquêtes. C'est également par enquête qu'on obtiendra certaines informations, factuelles ou non, qui nous aiderons à interpréter ce qui aura été observé.

4. Si nous opposons maintenant l'enquête non plus à l'observation mais à l'expérimentation, elle apparaît nécessaire chaque fois que ce qui nous intéresse ne relève pas de mécanismes de portée générale (par exemple : « Quelle est l'influence de la crédibilité de la source, ou l'influence exercée par un message ? »), mais de phénomènes tels qu'ils se produisent à un moment donné, dans une société donnée, avec leur complexité (par exemple : « Quelle est l'influence de tel journal sur ses différentes catégories de lecteurs ? »). Dans ce cas, la simplification du laboratoire, si elle permet de mieux contrôler ce qui se passe et de savoir à quels processus attribuer les phénomènes observés, ne permet pas une transposition immédiate à une situation réelle.

5. Enfin, et ce point est très proche du précédent, c'est par enquête, et parfois par observation, mais non par expérimentation, qu'on peut aboutir à des informations sur ce qui se passe à un moment donné. Quelle est la proportion d'électeurs qui a l'intention de voter pour tel parti ? Qui sont ces électeurs ? Comment la forme d'une ville affecte-t-elle les relations sociales ? Quels sont les processus qui amènent à tel engagement politique particulier ? L'expérimentation ne permet pas, ou du moins pas seule, de répondre à des questions de ce type, et l'observation, lorsqu'elle est possible, n'est, le plus souvent, pas suffisante, pour les raisons que nous avons examinées plus haut.

Il faut également mentionner, bien que ce type de justification se situe à un autre niveau que les précédents, que l'enquête est une technique relativement simple à appliquer. Même un entretien en profondeur ne dépasse que rarement deux heures, la passation d'un questionnaire peut se dérouler à peu près n'importe où, il n'y a pas besoin d'appareils compliqués, à l'exception parfois d'un magnétophone, ni de compères comme dans beaucoup d'expériences de psychologie sociale. Il en résulte que c'est à peu près la seule méthode qu'on peut, s'il le faut, appliquer à grande échelle, en choisissant les sujets. Tous ceux qui ont cherché désespérément à amener dans leur laboratoire des sujets autres que des étudiants en psychologie (et heureusement, pour les chercheurs en psychologie tout au moins, le nombre de ces derniers ne décroît pas), ou qui

ont passé des heures à l'affût de comportements intéressants à observer, apprécieront ces avantages pratiques.

Ceci dit, il n'en reste pas moins vrai que l'enquête reste totalement tributaire du langage. Tout ce dont on dispose, c'est de ce que le sujet a pu ou a voulu dire. On peut penser qu'une part importante de la méthodologie des enquêtes, et par conséquent de ce dont il sera question dans ce livre, vise à pallier, de diverses manières, cet inconvénient majeur, dans l'état actuel des connaissances linguistiques notamment, ceci ne signifiant nullement que la découverte d'une théorie sémantique ne lèverait pas cet obstacle. Mais cette pauvre théorie, promenée de la linguistique à la psychologie, n'est encore qu'une chimère enfançonne.

Il en résulte que, comme on le verra dans tout ce livre, et surtout dans les parties les plus liées au langage, comme celles qui portent sur les différents modes d'interrogation et sur l'analyse de contenu, on en est le plus souvent réduit à une *méthodologie négative*, indiquant les obstacles à éviter, les pièges où ne pas tomber, mais restant vague sur ce qu'il faut faire. Or, dans les cas où l'on peut indiquer des méthodes positives précises, leur intérêt ne réside pas tellement dans leur valeur intrinsèque, qui reste mal définie, que dans le caractère explicite des démarches proposées, ce qui constitue un avantage certain par rapport au flou artistique des méthodes interprétatives ou compréhensives, qui donnent des résultats certes séduisants, mais dont la validité reste incontrôlable.

Dans de telles situations, théoriquement et méthodologiquement incertaines, il faut se garder de deux dangers opposés. Le premier consiste à croire qu'une conclusion est valide parce qu'elle est intéressante, riche de perspectives, suggestive, vraisemblable. Le second, à n'admettre une conclusion comme valide que si elle a été obtenue par une méthode rigoureusement explicite et objective. Nous louvoyons constamment entre le Charybde du purisme méthodologique et le Scylla du laxisme intuitif et « concret ». Ici, nous avons choisi de privilégier le rigoureux et l'objectif, sans en ignorer toutes les limites. Toutefois, notre propos n'est pas dans une fuite en avant méthodologique qui masquerait les problèmes fondamentaux sans les résoudre, substituant ainsi un arsenal de méthodes aux validations scientifiques de l'adéquation entre le technique et le théorique. Mais agissant en fonction de critères d'utilité et posant que l'enquête permet une connaissance des événements qui peut ensuite donner lieu à des actions en milieu naturel, nous tenons à rendre cette connaissance la moins suspecte possible.

L'affinement méthodologique, dans ce cas, revient à chercher à réduire le taux d'erreur possible ou, plus précisément, à fournir au chercheur un recueil et une lecture des informations qui limite l'arbitraire et/ou les naïvetés du bon vouloir ou du bon sens.

Avant d'examiner de façon plus précise le champ d'application des enquêtes, et ses limites, il est nécessaire, pour éviter des malentendus fréquents, de distinguer la *méthode des enquêtes*, que nous entendons présenter ici dans sa généralité et sa complexité, et les *sondages d'opinion*, qui en sont une application. La distinction est importante car, très souvent, des critiques qui prétendent viser les enquêtes en général ne portent, en fait, que sur les seuls sondages d'opinion, qui constituent, et de loin, le mode d'enquête le plus connu en dehors des milieux tout à fait spécialisés.

Lorsque nous parlons d'enquête en général, nous entendons, comme nous l'avons

dit, toutes les formes d'interrogation d'individus à des fins de généralisation. Les sondages d'opinion en sont un cas particulier, tout d'abord par leur objectif : décrire l'état de l'opinion à un moment donné (l'opinion étant ici définie par les réponses à quelques questions très simples). Les conditions pratiques de réalisation de ces enquêtes imposent, d'autre part, un certain nombre de contraintes : les questionnaires doivent souvent être conçus très rapidement (par exemple, lorsqu'il s'agit de saisir les réactions à un événement inattendu, on ne peut évidemment pas prendre le temps qu'exigent toutes les étapes habituelles de préparation d'un questionnaire), les questions doivent être peu nombreuses si l'on veut trouver facilement des sujets et que l'enquête soit terminée rapidement ; elles doivent aussi être simples non seulement pour que toutes les personnes interrogées les comprennent facilement, ce qui est vrai pour toutes les formes d'enquêtes, mais aussi pour que leur sens et leur portée soient immédiatement apparents pour les lecteurs non spécialistes. D'autre part, au moins en ce qui concerne les résultats publiés dans la presse, les analyses des réponses recueillies sont réduites à quelques traitements très simples. En général, on se contente de donner les pourcentages des différentes réponses à chaque question, éventuellement la répartition de ces réponses en fonction de variables simples telles que l'âge, le sexe, la catégorie socio-professionnelle ou la préférence politique déclarée. On ne trouve jamais, par exemple, dans ces comptes rendus, l'étude simultanée des différentes réponses données par une même personne, ce qui permet d'apprécier leur cohérence, de dégager des types, etc. Cette simplicité dans la conception du questionnaire et dans l'analyse des réponses n'implique d'ailleurs pas que les résultats obtenus soient sans intérêt, ni que la réalisation d'un bon sondage d'opinion soit chose facile ; les risques d'erreurs ou de maladresses dans la formulation des questions sont nombreux, la représentativité de l'échantillon interrogé doit être assurée de façon particulièrement rigoureuse, plus rigoureuse même que pour d'autres types d'enquêtes plus ambitieuses par ailleurs.

Mais nous voudrions montrer que la méthode des enquêtes offre bien d'autres possibilités : en posant un plus grand nombre de questions, on peut faire des analyses plus approfondies, décrire de façon plus fine les opinions et les comportements qu'on cherche à étudier. On peut vérifier des hypothèses complexes. Par le recours à des modes de questionnement autres que le questionnaire, comme par exemple les entretiens non directifs, on peut saisir des aspects plus complexes des représentations et des attitudes.

Sans chercher à multiplier les exemples, essayons de voir, sur un cas, ce que peut apporter une enquête et quelles en sont les limites.

Après une élection, on veut connaître les déterminants des différents choix. On dispose, dans ce cas, de « tracés » qui ont l'avantage, rare, d'être en principe sans biais : les résultats électoraux et le taux de participation par bureau de vote ou par circonscription, les résultats aux élections précédentes ; d'autre part, le recensement donne, sur les mêmes unités, des renseignements socio-démographiques relatifs à la composition de la population concernée. Cela permet de procéder à une « analyse écologique » dans laquelle les unités prises en considération ne sont pas des individus, mais des groupes (dans le cas particulier, des circonscriptions). On pourra ainsi calculer la corrélation entre, par exemple, le pourcentage d'ouvriers dans la circonscription et la proportion des voix aux différents partis, comparer cette corrélation avec ce qu'elle était aux élections

précédentes, etc. De telles analyses sont possibles et fournissent des résultats certainement intéressants. Mais les données du recensement sont limitées : il y manque des facteurs qu'un sociologue pourra juger nécessaires, étant donné ses hypothèses, comme l'origine sociale, le type d'entreprise dans lequel l'électeur travaille, la religion, éventuellement des attitudes politiques ou plus générales. D'autre part, le passage des corrélations écologiques aux processus individuels pose des problèmes, et la transposition est délicate. Ce n'est pas parce que le pourcentage des voix à un certain parti est en corrélation avec la population d'ouvriers dans chaque circonscription qu'on peut en conclure immédiatement que les ouvriers votent de préférence pour ce parti. D'autres interprétations sont possibles.

Par rapport à l'analyse des statistiques existantes, l'enquête présente, dans ce cas, deux avantages : elle permet d'observer des relations au niveau des individus, et d'obtenir des informations plus riches sur chacun.

Mais l'enquête présente une limite très importante : la véracité des réponses peut être mise en doute. Des enquêtes effectuées dans des conditions permettant une validation des réponses par recoupement ont permis de montrer que les « erreurs » (volontaires ou non) sur le fait de déclarer avoir voté ou non, pouvaient être importantes et dépasser le tiers des réponses. Les mêmes biais peuvent fausser les réponses relatives au candidat pour lequel on a voté.

D'autre part, une enquête ne peut se faire, pour des raisons pratiques, que sur des échantillons restreints, ce qui peut être gênant, dans ce cas, étant donné l'importance des particularités locales. L'analyse écologique pourra, sans difficulté, être exhaustive.

Le choix final de la méthode, ou la décision d'utiliser les deux complémentairement, dépendra donc de la nature des hypothèses précises que l'on cherche à vérifier, et aussi des moyens dont on dispose.

1.3. CONCLUSION

L'étude de l'enquête, considérée comme objet d'analyse, rencontre plusieurs difficultés dont deux sont - à notre sens - essentielles.

La première réside dans le fait que les techniques d'enquête sont considérées dans la plupart des cas comme de pures méthodologies et qu'elles ne peuvent à ce titre que difficilement être questionnées sur leurs fondements théoriques et épistémologiques. En effet, les techniques d'enquête sont reliées dans tous les cas à un support « langagier ». Ce qui est à considérer, c'est donc une théorie (ou une absence de théorie) du langage et plus précisément du sens. La question sémantique paraît primordiale si l'on veut éviter les pièges issus des possibilités de lecture multiple d'un même message.

Les techniques d'enquête, *en tant que telles*, ne peuvent pas fournir de réponse à ce problème précis ; tout au plus peuvent-elles limiter les risques de lecture multiple et systématiser les approches.

Ce qui fonde les méthodes d'enquête ne peut donc fournir de réponse *directe* aux difficultés qui surgissent de l'objet (le langage) auquel elles s'appliquent. Les risques d'ignorer les problèmes théoriques du sens et de centrer son attention sur le traitement

des données sont accrus par cette possible fuite en avant méthodologique, d'autant plus sécurisante qu'elle est porteuse de réponses, apparemment du moins.

La seconde est liée à la double appartenance logique des techniques d'enquête. A la fois logique-procès et logique-formalisée, elle est soumise à une double intention.

Le caractère mathématique du traitement de l'information est corrélatif d'un désir d'univocité dans la transmission et la lecture. Les risques d'assimilation abusive de l'univocité du langage de l'outil aux présupposés qui ont permis sa construction d'une part (théorie des attitudes par exemple, dans le cas des échelles d'attitude), à l'objet auquel il s'applique (le langage) d'autre part, sont grands. Et si une analyse critique est à même d'éviter ces biais, la pratique courante est encline à les ignorer. Le recueil de l'information (que ce soit à travers des questionnaires ou des interviews), qui est également construction progressive de l'objet, participe de ce fait à une logique-procès. On entend par-là que poser une question et partant un ensemble de questions, c'est progressivement limiter — par une procédure d'entonnoir — le champ des possibles concernant un objet quelconque. Agir ainsi revient en fait à constituer progressivement un objet, mais sans garantir l'univocité de la lecture.

Ces deux difficultés conduisent à nettement distinguer deux préoccupations : l'une scientifique, l'autre méthodologique.

Ignorer que la distinction existe et confondre par là même méthodologie et science, c'est s'exposer à ne pas poser de questions sur ce qui fonde telle ou telle pratique.

Dans toute recherche, la première étape consiste à préciser le problème, à fixer les objectifs et à formuler les hypothèses qu'on cherche à vérifier. Malgré son importance, nous ne dirons rien ici de cette phase, d'une part parce que le problème déborde largement notre sujet, qui est l'exposé et la discussion d'une méthode, et d'autre part, parce qu'il n'y a pas grand-chose à en dire. On le sait bien, dans les sciences sociales, les idées, les hypothèses viennent d'où elles peuvent : il n'y a aucune règle à ce propos, tous les moyens sont bons. Elles peuvent être déduites rigoureusement d'une théorie, provenir d'un problème pratique ou d'un étonnement devant tel aspect de la vie quotidienne, peu importe. Tout l'effort de rigueur scientifique porte sur les méthodes à mettre en œuvre une fois le problème posé. Ceci est d'ailleurs significatif de l'état actuel de développement des sciences sociales. En physique, les problèmes, les hypothèses proviennent d'un corps antérieur de connaissances et de théories, au moins dans le cadre de ce que T.S. Kuhn appelle la « science normale », c'est-à-dire celle qui se développe en dehors des périodes de crise, où des difficultés les amènent souvent à remettre en cause des parties importantes de l'acquis.

Une fois les hypothèses posées, il reste à choisir la méthode, ou souvent les méthodes, qu'on emploiera. Des contraintes institutionnelles, sociales, idéologiques pèsent sur ce choix, indépendamment des exigences proprement scientifiques. Quelle que puisse être la conviction intime du chercheur, il lui sera le plus souvent difficile de convaincre un client soucieux d'applications pratiques immédiates qu'une expérience en laboratoire pourra l'éclairer. Inversement, dans un cadre universitaire, une expérience bien montée et ingénieuse a plus de chances d'apporter du prestige à son auteur qu'une enquête ou un ensemble d'observations. Ces dernières, de plus, soulèvent des problèmes déontologiques évidents. Enfin, les différentes méthodes possibles sont plus ou moins

coûteuses. Ce dernier point est particulièrement important en ce qui concerne les enquêtes, qui exigent d'interroger un grand nombre de sujets, et des dépouillements longs et complexes. Dans la discussion ultérieure des techniques, c'est un point à garder présent à l'esprit : bien des raffinements des méthodes d'échantillonnage, par exemple, se justifient en premier lieu par le souci d'obtenir le maximun d'information au moindre coût.

Ainsi les techniques d'enquête renvoient à une double rationalité méthodologique et pratique. Vouloir oublier l'une ou l'autre condamne nécessairement à commettre des erreurs.

Dans ce livre, nous avons voulu nous adresser à celui que nous avons appelé le chercheur, et non au statisticien ou à l'informaticien. Cela marque les limites de l'ouvrage. En particulier, nous avons systématiquement renoncé à présenter, autrement que par des allusions, certaines techniques dites « de pointe », comme l'analyse automatique du discours, l'analyse des correspondances, ou l'analyse algébrique des questionnaires. Leur exposé, et surtout leur discussion, aurait exigé de longs développements techniques, et exagérément allongé le texte. De plus, nous préférons mettre en garde contre une fuite en avant méthodologique qu'on observe actuellement chez certains spécialistes : maîtrisant déjà mal les méthodes élémentaires, ils croient être plus scientifiques en recourant à des techniques sophistiquées, dont ils connaissent mal les hypothèses et les limites. Le recours à l'ordinateur, outre la garantie de scientificité qu'il assure aux yeux de certains, permet, par son automatisme, d'obtenir des résultats complexes sans effort, mais aussi sans aucun contact avec les données elles-mêmes, contact qu'on abandonne volontiers aux « codeurs ». Nous pensons que, entre les résultats sommaires des sondages d'opinion et l'emploi de méthodes lourdes, il y a place pour des procédures intermédiaires, dont il est nécessaire, et parfaitement possible même quand on n'est ni mathématicien, ni linguiste, de comprendre la *logique*.

Les différentes étapes d'une enquête ne se déroulent pas selon un ordre linéaire et constant : si, logiquement, on peut penser qu'il faut commencer par savoir qui on veut interroger, puis ce qu'on veut demander, avant d'analyser les réponses, l'ordre réel est beaucoup moins net. C'est le type d'analyse prévu qui oriente vers telle méthode d'interrogation ou telle autre, qui guide la conception du questionnaire. A son tour, la méthode d'interrogation déterminera le choix de l'échantillon. C'est pourquoi il ne faut pas prendre trop à la lettre le plan que nous avons adopté, et qui suit l'ordre « logique ».

BIBLIOGRAPHIE

Il n'existe pas d'ouvrage d'ensemble, en français, sur les enquêtes. En anglais, on pourra lire :

- C.A. Moser, G. Kalton, *Survey Methods in Social Investigations*, 2nd ed. Heineman, London, 1971, qui contient de très nombreuses indications pratiques.

Pour situer les enquêtes parmi les méthodes de la sociologie et de la psychologie sociale :

- Préface de S. Moscovici à D. Jodelet, J. Viet, P. Besnard, *La Psychologie sociale, une discipline en mouvement*, Mouton, Paris-La Haye, 1972.
- P. Bourdieu, J.P. Passeron, J.C. Chamboredon, *Le Métier de sociologue*, Mouton, Paris-La Haye, 1968.
- J. Grisez, *Méthodes de la psychologie sociale*, Paris, P.U.F., 1975.
- A. Cicourel, *Method and Measurement in Sociology*, The Free Press, New York, 1964. Discussion très critique des méthodes de la sociologie positive.

Sur les sondages d'opinion :

- J. Stœtzel, A. Girard, *Les sondages d'opinion publique*, Paris, P.U.F., 1973.
- F. Bon, *Les sondages peuvent-ils se tromper ?*, Paris, Calmann-Lévy, 1972. Présentation informée et claire, à l'usage du grand public, des principaux aspects des sondages d'opinion.

Première partie

LA PRODUCTION DES DONNÉES D'ENQUÊTES

Chapitre 2

QUI INTERROGER ?

Une fois le problème défini, et les hypothèses posées, lorsqu'on s'est assuré que c'est bien l'enquête, venant éventuellement en complément d'autres techniques, qui constitue la méthode la mieux adaptée, et une fois ses objectifs spécifiques explicités, la première question à se poser est : « Qui interroger ? », question qui elle-même se décompose en deux autres :

a) Quelle est la population (ou peut-être les populations) qu'il est nécessaire de connaître ?

b) Comment choisir, dans cette population, les personnes qu'on interrogera effectivement, étant donné que, dans la plupart des cas, il est exclu, et d'ailleurs inutile, de les interroger toutes ?

2.1. LE CHOIX DE LA POPULATION A INTERROGER

La première question n'est pas toujours explicitée : on s'attache plus souvent à la seconde, qui recouvre les problèmes des méthodes d'échantillonnage et de taille de l'échantillon, problèmes auxquels les statisticiens ont apporté des solutions précises. En revanche, le choix de la population à étudier est beaucoup moins formalisé, ce qui amène le plus souvent à passer sous silence la première question. On lui apporte bien

une réponse, par la force des choses, mais souvent sans y réfléchir vraiment, comme si l'option prise allait de soi, était la seule possible. Or, c'est très rarement le cas, et il est souvent possible d'améliorer considérablement la portée d'une enquête, ou d'en diminuer le coût, en choisissant de façon adéquate les populations à prendre en considération.

Prenons d'abord un cas simple, celui des sondages pré-électoraux : c'est un exemple où le problème est immédiatement résolu : on s'intéressera à la totalité des électeurs, et à eux seulement. Mais de tels cas, où la population concernée est définie explicitement par la loi, sont extrêmement rares. Toutefois, si l'on veut élargir le problème et tenter de comprendre une dynamique politique plus générale, ou si l'on veut étudier « l'opinion publique », il est plus difficile, moins évident, de définir la population intéressée : ne prendre en considération que les seuls électeurs, comme dans l'exemple précédent, est maintenant trop restrictif : les étrangers, les jeunes, ceux qui, pour des raisons diverses, ne sont pas inscrits sur les listes électorales font, eux aussi, partie de « l'opinion publique », et les exclure a priori constitue déjà un choix politique ou idéologique [1].

Mais, si on décide de les intégrer dans la population étudiée, il faut bien en définir les limites. Par exemple, on pourra n'y intégrer que les étrangers établis en France depuis un certain temps, excluant ainsi non seulement les touristes, mais aussi les immigrés trop récemment arrivés ; le choix de la durée pose évidemment des problèmes. Il en va de même pour les jeunes : à partir de quel âge va-t-on les interroger ? Cela dépend de l'image que l'on a de l'évolution des adolescents et de leur maturité, et du problème que l'on traite. Dans chaque cas, il est nécessaire de faire des choix, de prendre des décisions, de placer des limites partiellement arbitraires, même si, lorsque les résultats du sondage seront publiés, ils le seront simplement sous la forme : « Les Français pensent que... »

Dans cet exemple, on cherche en principe à interroger « tout le monde » [2], même si, finalement, il y a nécessairement des exclus, ne serait-ce que les petits enfants. Dans d'autres cas, c'est à un groupe plus restreint qu'on peut s'intéresser : l'objet de la recherche pourra être les paysans, les cadres, les jeunes, les habitants d'un quartier, etc. Ceci suppose qu'on ait clairement défini ce qu'on entend par-là, c'est-à-dire qu'on soit en mesure, devant tout individu, de dire s'il fait partie ou non de la population concernée. S'il s'agit de jeunes, il faudra fixer les âges minimum et maximum ; si l'enquête doit porter sur les cadres, une définition précise devra être donnée à l'avance, indiquant si on retient les cadres moyens, certains techniciens, ou les directeurs généraux, etc.

Très souvent, le chercheur qui s'intéresse à une population particulière peut être amené à interroger en plus une autre population, pour disposer d'un terme de compa-

---

1. Ce peut être aussi parfois un simple choix pratique, les listes électorales constituant, nous le verrons plus loin à propos des méthodes d'échantillonnage, une base de sondage très commode. Mais ce qui nous intéresse ici, c'est la population *visée*. Nous verrons au chapitre suivant les distorsions qu'on est parfois amené à accepter pour des raisons pratiques.

2. Décider de prendre en considération, pour une étude d'opinion publique, l'ensemble de la population constitue déjà un choix important. C'est se représenter l'opinion sur le modèle du suffrage universel qui donne à chacun le même poids, en ignorant les processus réels qui constituent cette opinion. Lorsque les plus prudents parmi les spécialistes des sondages d'opinion mettent en garde leurs lecteurs contre les extrapolations trop larges de leurs résultats, affirmant qu'ils ne donnent que l'image de l'opinion à un moment donné, c'est à cette absence de prise en considération des aspects dynamiques qu'ils font, entre autres, allusion. Si l'on voulait étudier l'évolution de l'opinion, il faudrait constituer des échantillons plus complexes, où, entre autres, chacun n'aurait pas le même poids, et ne serait pas interrogé de la même manière que tous les autres.

raison. Si, par exemple, son travail porte sur les jeunes, il lui sera peut-être nécessaire d'interroger également une population plus âgée, pour pouvoir déterminer sans ambiguïté ce qui, dans les réponses données par les jeunes, leur est vraiment spécifique. Certes, on peut se contenter de n'interroger que des jeunes, et ne présenter les résultats de l'enquête que comme une observation de ce que pensent, ou font, les jeunes. Mais de tels résultats sont presque inévitablement interprétés par comparaison avec une représentation antérieure de ce que sont les jeunes ou les adultes. Parmi toutes les réponses obtenues, certaines seront considérées comme spécifiquement « jeunes » alors qu'on admettra que d'autres devraient se retrouver identiques dans une population plus large. Le seul moyen d'éviter ces a priori, c'est le recours à un « groupe contrôle ». La composition de ce groupe n'est pas toujours évidente : dans notre exemple, qu'est-ce qu'un groupe de « non-jeunes » ? Tous ceux dont l'âge est supérieur à la limite prévue ? Ou n'est-il pas préférable de ne retenir qu'une population plus homogène, par exemple d'âge compris entre 30 et 40 ans ? Quelle que soit la solution adoptée, on se posera peut-être une question complémentaire : à quelles conditions un échantillon « jeune » et un échantillon « âgé » sont-ils comparables, c'est-à-dire ne différant que par l'âge ? Il est évident que, si on ne prend aucune précaution particulière, les deux échantillons différeront par toute une série d'autres caractéristiques telles que la proportion de célibataires, l'insertion professionnelle, le revenu, le nombre d'enfants, le niveau d'instruction (on trouvera une plus forte proportion de primaires parmi les plus âgés), etc. Toutes ces caractéristiques sont effectivement liées à l'âge, et on peut penser que la différence entre les jeunes et les plus agés n'est rien d'autre que le résultat de toutes ces différences. Dans ce cas, on comparera globalement les deux populations. Mais si on veut aller au-delà d'une simple description, si on veut interpréter les différences et comprendre ce qui se passe lorsqu'on avance en âge, il faudra peut-être utiliser des « échantillons expérimentaux », plus artificiels, mais comparables du point de vue de toutes les variables qu'on juge pertinentes, ce qui suppose qu'on ait des hypothèses à leur sujet.

Dans tous ces exemples la population intéressante est définie au départ, même s'il reste nécessaire d'en préciser les limites. Mais il y a aussi des cas où le choix de la population à étudier n'est pas évident. Dans une enquête commerciale, par exemple, l'objectif pourra être d'étudier les motivations et les freins à l'achat d'une certaine marque d'un produit, en vue de la définition d'une politique publicitaire. Pour cela suffit-il d'interroger seulement les acheteurs de cette marque ? Tous les acheteurs du produit en question ? Ou élargira-t-on la population à étudier aux non-acheteurs, essayant ainsi d'atteindre la clientèle potentielle ? S'il s'agit d'un produit dont on sait qu'il est acheté le plus souvent par les femmes, comme un produit alimentaire, va-t-on se contenter de n'interroger qu'elles, ou s'intéressera-t-on aussi aux réactions des hommes et des enfants qui, s'ils ne sont que rarement acheteurs eux-mêmes, n'en sont pas moins consommateurs et donc peuvent influencer l'achat ? Bien souvent, la réponse à de telles questions ne peut être apportée que par une étude exploratoire préalable, nécessairement très informelle puisqu'on ne dispose que d'hypothèses encore vagues. Mais ceci permettra de délimiter le problème et de se donner les moyens d'effectuer de façon moins arbitraire certains des choix dont nous avons montré la nécessité.

Une fois la population à étudier ainsi définie, et les problèmes que nous venons d'évoquer résolus, on peut se rendre compte que certaines analyses, certaines comparaisons, nécessaires pour atteindre les buts poursuivis, risquent de n'être pratiquement pas possibles, certains sous-groupes intéressants n'étant pas suffisamment représentés dans la population retenue, et donc dans l'échantillon, pour pouvoir être utilisés. C'est souvent le cas, par exemple, des cadres, qui sont suffisamment importants du point de vue sociologique et économique pour qu'on souhaite pouvoir en parler en détail, et les comparer aux autres catégories sociales. Mais un échantillon représentatif de taille raisonnable de l'ensemble de la population n'en contiendra qu'un nombre trop faible pour permettre des analyses un peu fines. Augmenter la taille de l'échantillon pour que le nombre de cadres soit suffisant est coûteux, et amène à inclure un nombre inutilement grand d'individus d'autres catégories. On pourra plutôt décider de constituer un échantillon où les cadres seront surreprésentés par rapport à leur représentation réelle dans la population. Cette manière de faire permettra d'analyser de façon approfondie les réponses des cadres et de les comparer à celles des autres groupes, mais on ne peut plus dire, en toute rigueur, que c'est la population totale qui est l'objet de l'enquête. Il n'est même plus légitime de parler d'une « population parente » d'un tel échantillon. En fait, il s'agit de deux échantillons distincts, l'un de cadres, l'autre des autres catégories sociales, et en principe on ne devrait rien dire sur leur union ; tout ce qu'on peut faire, c'est les analyser séparément et les comparer. Il n'est pas possible d'en tirer directement des inférences sur la population totale. Toutefois, si effectivement toutes les *estimations* sont impossibles (on ne peut pas, à partir d'un tel échantillon composite, énoncer des propositions telles que « n % de la population dit que... »), en revanche certaines *relations* entre variables ne sont pas affectées, et peuvent donc être étudiées même dans ces conditions. Si certaines estimations sont, malgré tout, jugées nécessaires, il est toujours possible de tirer de l'échantillon des cadres un sous-échantillon de la taille voulue pour reconstituer un échantillon représentatif de la population totale. Ou, pour éviter une perte d'information, calculer la valeur cherchée en pondérant chaque sous-échantillon en fonction inverse de son taux de sondage (cf. 2.2.1.4.).

Lorsque nous avons parlé de *groupe contrôle*, nous avons repris une notion de la méthodologie expérimentale ; de même, lorsque nous avons mentionné la possibilité de construire un échantillon en fonction des analyses qu'on veut effectuer, et non en fonction de la seule représentativité, nous avons eu recours à un des procédés fondamentaux de l'expérimentation. En fait, les transpositions de l'*attitude expérimentale* à la méthodologie des enquêtes peut être poussée très loin ; on peut parfaitement construire des échantillons en fonction de *plans d'expérience* aussi complexes que ceux qui sont utilisés en laboratoire. Plutôt que de constituer un échantillon représentatif d'une population, on construit une série d'échantillons représentatifs de sous-populations, la taille et les caractéristiques de chacun de ces échantillons étant déterminées par les analyses qu'on a l'intention d'effectuer. Mais cela suppose que, comme au laboratoire, on ne cherche à vérifier qu'un petit nombre d'hypothèses très précises, en fonction desquelles on construira le plan expérimental. Cependant, lorsqu'on sort du laboratoire, il devient beaucoup plus difficile de maîtriser tous les facteurs ; on est obligé de tenir compte d'un très grand nombre d'entre eux. Lorsqu'on dispose d'un échantillon

représentatif de taille suffisante, il est possible de vérifier un grand nombre d'hypothèses, d'étudier l'influence d'un grand nombre de facteurs, mais souvent de façon très incertaine, à cause du trop faible nombre de représentants de certains groupes. Le plan d'expérience, lui, permet de vérifier un petit nombre d'hypothèses précises, avec peu de sujets, mais il est peu souple, et vérifier de nouvelles hypothèses devient souvent très difficile. Nous reviendrons sur ces divers problèmes lorsque nous parlerons des méthodes d'analyse des résultats ; c'est elles, évidemment, qui conditionnent le choix de la méthode à adopter.

2.2. LES MÉTHODES D'ÉCHANTILLONNAGE

2.2.1. Les échantillons représentatifs

2.2.1.1. Définition

Il est très rare qu'on puisse étudier exhaustivement une population, c'est-à-dire en interroger tous les membres : ce serait si long et si coûteux que c'est pratiquement impossible. D'ailleurs c'est inutile : interroger un nombre restreint de personnes, à condition qu'elles aient été correctement choisies, peut apporter autant d'information, à une certaine erreur près, erreur calculable, et qu'on peut rendre suffisamment faible. Le problème est de choisir un groupe d'individus, un *échantillon*, tel que les observations qu'on fera sur lui pourront être généralisées à l'ensemble de la population ; il faut donc que l'échantillon présente les mêmes caractéristiques que la population, qu'il soit *représentatif*.

En principe, un échantillon est représentatif si les unités qui le constituent ont été choisies par un procédé tel que *tous les membres de la population ont la même probabilité de faire partie de* l'échantillon. Si ce n'est pas le cas, on dira que l'échantillon est *biaisé* : puisque certains individus avaient plus de chances que d'autres d'être choisis, les catégories auxquelles ils appartiennent occuperont dans l'échantillon plus de place qu'elles ne le devraient ; les caractéristiques de l'échantillon seront donc systématiquement différentes de celles de la population.

Essayons de préciser ces notions par quelques exemples qui montrent ce qu'implique cette définition, et les difficultés qu'on rencontre, pratiquement, à réaliser de tels échantillons.

Supposons que nous voulions étudier l'ensemble de la population française. On ne peut évidemment pas se contenter d'interroger « n'importe qui », ses amis, la concierge, et son marchand de journaux. Descendre dans la rue et accoster les passants « au hasard » exclut certainement moins de monde, élargit l'étendue de l'échantillon possible, mais encore faut-il prendre garde à se placer, pour cela, dans un nombre suffisant d'endroits différents. D'autre part, même sans mentionner pour le moment les biais introduits par ceux qui refusent de répondre, il ne faut pas oublier qu'une personne a une probabilité d'autant plus élevée de faire partie de l'échantillon qu'elle passe plus de temps dans la rue. Ce pourra être parce qu'elle aime flâner et qu'elle a le temps de le

faire, ou que son travail ou ses autres activités l'amènent à sortir beaucoup. Ceux qui se déplacent en voiture auront moins de chances d'être touchés que les piétons. Les personnes âgées, les malades, risquent d'être pratiquement exclus. On voit que la condition qui définit un échantillon représentatif est loin d'être satisfaite. Néanmoins, au prix d'un certain nombre de précautions, sur lesquelles nous reviendrons en détail, la méthode peut être utilisable dans certains cas. Elle est, en tout cas, largement utilisée.

Si maintenant nous disposons d'une liste de toutes les personnes habitant la France, par exemple si nous avons accès aux fiches du recensement (ce qui, précisons-le tout de suite, n'est presque jamais possible), nous pouvons y tirer au sort un nombre égal (en pratique supérieur, pour tenir compte des absences et des refus possibles) à la taille voulue de l'échantillon. Dans ce cas idéal, on disposerait effectivement d'un véritable échantillon représentatif, pouvant satisfaire le plus rigoriste des statisticiens : toutes les personnes habitant en France ont effectivement la même probabilité d'être choisies.

Mais, cette façon de procéder, parfaitement rigoureuse, est bien souvent pratiquement inapplicable dans tous ses détails. En effet, la première condition, c'est de disposer d'une liste exhaustive des membres de la population à étudier, ce qu'on appelle une *base de sondage*. Or de telles listes sont rares, et encore plus rarement accessibles. Ensuite, en supposant qu'elle existe et qu'on y ait accès, l'échantillon sera trop dispersé géographiquement, puisque des individus résidant en n'importe quel point du territoire peuvent être tirés. Les enquêteurs perdront beaucoup trop de temps en voyages qui pourront parfois être très longs pour toucher une ou deux personnes seulement. Il est donc nécessaire de trouver d'autres techniques.

Passons maintenant à un exemple opposé : ce n'est plus l'ensemble de la population qui nous intéresse, mais un groupe très restreint, disons les acheteurs d'un certain produit peu courant. Dans ce cas, il est évident qu'il n'existe pas de liste exhaustive, donc pas de base de sondage au sens strict. On peut envisager de constituer un échantillon représentatif de l'ensemble de la population, d'en contacter tous les membres, et de ne retenir pour l'enquête que les consommateurs du produit en question. Mais, pour constituer ainsi un échantillon final de taille suffisante, le premier devra être gigantesque. Une autre manière de procéder pourra consister à se poster là où le produit est vendu, et à accoster tous ceux qu'on verra l'acheter. Mais, de cette manière, les acheteurs fréquents, que ce soit parce qu'ils sont gros consommateurs ou parce qu'ils achètent peu à la fois, auront une probabilité plus forte que les autres d'être touchés. De plus, les caractéristiques de l'échantillon obtenu dépendent du choix des points de vente.

Ces exemples nous semblent soulever l'essentiel des problèmes pratiques que pose la constitution d'un échantillon satisfaisant : un échantillon représentatif au sens strict étant le plus souvent impossible à obtenir, que peut-on faire pour avoir malgré tout un échantillon acceptable et adapté aux besoins ? Le développement des techniques de sondage vise à apporter des réponses à cette question.

2.2.1.2. La constitution des échantillons aléatoires

Les échantillons dits aléatoires, ou statistiques, s'obtiennent par un tirage au sort respectant la condition de définition des échantillons représentatifs : faire en sorte que

chaque membre de la population ait la même probabilité de faire partie de l'échantillon. Pour cela, la situation idéale est celle où l'on dispose d'une liste exhaustive de la population, la *base de sondage*.

Il n'existe actuellement en France aucune base de sondage *accessible* couvrant la totalité de la population [1]. Seul, l'INSEE, avec le recensement, dispose d'une telle base, ce qui leur permet ensuite de réaliser ses autres enquêtes sur des échantillons statistiques rigoureux, mais ces documents ne sont en principe ouverts qu'aux chercheurs de l'INSEE, pour les besoins de cet organisme. Les autres listes qui existent, comme par exemple celles des contributions directes ou le fichier des abonnés E.D.F., sont bien exhaustives, à peu de chose près, mais ne sont ouvertes en principe à personne.

Le chercheur en est donc réduit à des approximations. La meilleure, à notre connaissance, est constituée par les listes électorales, auxquelles on peut éventuellement avoir accès après avoir obtenu l'autorisation du Maire, autorisation qui est parfois accordée pour des enquêtes d'intérêt général, ou de recherche fondamentale. Mais ces listes sont loin de recouvrir la totalité de la population. Les étrangers et les mineurs en sont exclus par définition, et l'inscription des Français majeurs n'est pas obligatoire. Certaines personnes refusent de s'y inscrire par principe, d'autres, des jeunes surtout, négligent de le faire, laissant passer le moment de la révision des listes. D'autre part, bien des électeurs sont, pour des raisons diverses, inscrits dans une circonscription différente de celle de leur domicile réel : résidence secondaire, domicile antérieur, etc., ce qui les rend impossibles à atteindre à l'adresse indiquée. Malgré tous ces défauts, les listes électorales restent une base de sondage assez souvent employée, faute de mieux, bien que la généralisation à l'ensemble de la population ne soit, en toute rigueur, pas légitime. Mais on considère souvent des échantillons obtenus à partir de ces listes comme des approximations acceptables.

L'annuaire du téléphone, largement utilisé comme base de sondage aux Etats-Unis, est évidemment très incomplet en France, mais peut néanmoins être acceptable si l'on s'intéresse aux ménages ayant des revenus élevés. Lorsque l'échantillon aura été tiré et l'enquête effectuée, une comparaison entre les caractéristiques de cet échantillon (du point de vue de l'âge, du sexe, et de la catégorie socio-professionnelle, par exemple) avec les données statistiques globales dont on dispose par ailleurs permettra de préciser les écarts par rapport à la population totale.

Lorsqu'on s'intéresse à une population particulière, il arrive qu'il existe pour elle une base de sondage spécifique, et qu'on puisse y avoir accès. Pour obtenir un échantillon de salariés d'une entreprise, par exemple, les listes établies à la comptabilité pour la paie et les charges sociales doivent être exhaustives et exactes, et pourront être utilisées si la direction le veut bien.

L'opération pratique de tirage au sort ne doit pas se faire n'importe comment, sous prétexte qu'on procède de toute façon au hasard ; certaines précautions sont indispensables. L'ordre d'apparition des noms sur certaines listes peut parfois être considéré comme aléatoire, ce qui simplifie la procédure. C'est le cas, par exemple, de l'ordre al-

---

1. Sauf dans le Nord, où existe l'Annuaire Ravet-Anceau qui regroupe, en principe, la totalité de la population, donnant les noms, les adresses et les professions.

phabétique, qu'on peut considérer comme un ordre au hasard, mis à part le fait que, à la fin de l'alphabet, on trouve en France une plus grande proportion de noms d'origine étrangère. Mais d'autres listes peuvent être classées selon leur ordre d'arrivée dans la population en question, ou être regroupées en fonction de caractéristiques diverses.

Quelle que soit la structure de la liste, la meilleure manière de se prémunir contre tous les biais est de procéder à un véritable tirage aléatoire, en utilisant une table de nombres au hasard. On y prélève autant de nombres que d'unités souhaitées, et on va rechercher dans la liste les noms occupant les rangs correspondants.

Si l'on est certain que la liste dont on dispose est bien construite selon un ordre aléatoire, on peut procéder de façon plus simple : on se contente de tirer au sort un premier nom, par un procédé quelconque, puis, si le taux de sondage est 1/n, on prendra les noms ayant les rangs n, 2n, 3n..., après ce point de départ, puis on revient quand il le faut au début de la liste et on recommence.

Il faut toujours prendre garde de tirer plus de noms que ceux dont on a besoin, pour tenir compte des erreurs éventuelles, des personnes absentes et de celles qui refusent de répondre. Selon la population visée, ce supplément passera de 30 % au minimum, jusqu'à 100 % ou même plus.

Il arrive que la base de sondage dont on dispose soit trop large par rapport à la population visée. Ce sera le cas, par exemple, si l'on veut constituer un échantillon de salariés, et qu'on ne dispose que de listes électorales où les professions ne sont, en général, pas indiquées. On peut alors s'appuyer sur le fait que tout *sous-échantillon d'un échantillon représentatif est lui-même représentatif de sa sous-population*. Il suffira donc de tirer un échantillon dans la liste dont on dispose, et de demander à toutes les personnes ainsi désignées quel est leur statut professionnel, n'interrogeant que celles qui déclarent être salariées. Cela augmente le nombre de contacts nécessaires, donc le coût et la durée de l'enquête, mais l'échantillon ainsi obtenu sera aussi satisfaisant que la base de sondage elle-même : il aura les mêmes qualités et les mêmes biais.

Certaines catégories non représentées sur une base de sondage peuvent malgré cela être incluses dans un échantillon si on peut les rattacher à une des unités tirées. Par exemple, on peut constituer un échantillon correct de mineurs à partir de listes électorales, dont ils sont pourtant exclus par définition, en interrogeant tous les enfants des personnes majeures tirées. L'échantillon de jeunes ainsi obtenu sera représentatif si l'échantillon d'adultes pouvait être considéré comme tel.

Lorsque, par chance, on dispose d'une base de sondage comportant, outre le nom et l'adresse, d'autres indications, on est souvent tenté de vouloir « améliorer » l'échantillon en tirant séparément des sous-échantillons de chaque catégorie. Par exemple, si l'on fait un tirage au sort dans le fichier du personnel d'une grosse entreprise, certains pensent assurer une meilleure représentativité en tirant des échantillons distincts dans chaque service, par exemple, de manière qu'on retrouve dans l'échantillon total la proportion réelle des membres de chacun d'eux, craignant qu'un simple tirage au sort général ne donne pas une image suffisamment fidèle. En fait, ces précautions sont le plus souvent complètement inutiles, et même nuisibles. Un échantillonnage aléatoire véritable assure de lui-même la représentativité, même si l'on ne retrouve pas exactement les proportions réelles de chaque catégorie, les écarts resteront à l'intérieur de limites que la

théorie des probabilités permet de calculer, et dont les tests statistiques tiennent compte. En essayant d'imposer des contraintes à la structure de l'échantillon, on ne se trouve plus dans la situation supposée par ces hypothèses, qui ne sont plus applicables. Malgré les apparences, on maîtrise moins bien un échantillon ainsi « forcé » qu'un échantillon aléatoire, dont on connaît exactement le mode de construction.

Ce n'est que lorsque certaines catégories sont faiblement représentées dans la population que des précautions supplémentaires sont nécessaires. Nous l'avons vu en parlant de la détermination des populations à interroger : il peut être utile d'utiliser des taux de sondage différents par catégorie, pour éviter que certaines ne soient représentées que par un trop petit nombre de sujets pour qu'on puisse en tenir compte. De tels échantillons *stratifiés* posent toutefois des problèmes lors de l'estimation de certaines grandeurs : il importe alors de redresser l'échantillon, c'est-à-dire de redonner à chaque strate son poids réel, par exemple en multipliant les observations recueillies dans chaque strate par un coefficient inversement proportionnel au taux de sondage utilisé dans cette strate.

On voit que, malgré ses règles strictes, rigides, la constitution d'échantillons représentatifs par un procédé aléatoire est en fait plus simple qu'il n'apparaît au premier abord. La difficulté majeure réside en premier lieu dans la rareté des bases de sondage à la fois existantes et accessibles au chercheur ou au praticien. Il en résulte que, bien que la constitution d'échantillons aléatoires soit le procédé le plus rigoureux, il est rarement utilisé en pratique, et qu'on lui préfère des procédés beaucoup moins satisfaisants mais plus facilement et plus généralement utilisables (cf. 2.2.2. et 2.2.3.).

2.2.1.3. Unités et grappes

Le plus souvent, une enquête sociologique porte sur des *individus*, mais il peut arriver qu'on s'intéresse à des *ménages* [1], à des *entreprises*, à des *unités administratives*, à des *associations*, etc.

Par exemple, si l'on veut étudier les dépenses de consommation, l'équipement électroménager ou certains problèmes familiaux, c'est le ménage qui constituera l'unité d'analyse, et non l'individu.

Les bases de sondage disponibles, elles, ne portent pas nécessairement sur l'unité qui intéresse le chercheur. Par exemple, on veut interroger des individus, et on ne dispose que d'un fichier de logements ; on s'intéresse à des salariés et on a accès à une liste d'entreprises. Ou, inversement, on veut constituer un échantillon de ménages, et la seule base accessible est constituée par les listes électorales, qui regroupent des individus.

Il y a donc lieu de distinguer les *unités de sondage*, que l'on trouve dans les bases disponibles, et les *unités d'analyse* qui sont celles sur lesquelles portent les hypothèses, et qu'on désire toucher. Lorsque les deux ne coïncident pas, il faut pouvoir passer de l'une à l'autre.

---

1. Nous prenons ici ménage au sens de l'INSEE, désignant ainsi tout ensemble de personnes habitant le même logement.

Prenons comme exemple le cas, fréquent, où l'unité d'analyse est l'individu, mais où l'on ne dispose, comme base de sondage, que d'un fichier de logements, donc, à peu de chose près, de ménages. Si l'on n'interroge qu'une personne par ménage, quelle que soit la manière de la choisir, l'échantillon ne sera pas représentatif : une personne aura d'autant moins de chance de faire partie de l'échantillon qu'elle appartiendra à un ménage plus grand. En effet, si tous les ménages ont la même probabilité d'être tirés et donc de faire partie de l'échantillon, un individu vivant seul, et dont le logement a été tiré, a une probabilité de 1 de faire partie de l'échantillon, alors que celui qui appartient à un ménage de 4 personnes, n'aura qu'une probabilité de 1/4 d'être interrogé. Ce n'est donc pas une solution satisfaisante.

En revanche, en interrogeant systématiquement tous les membres du ménage entrant dans la population étudiée, on obtient un échantillon non biaisé d'individus. En effet, la probabilité d'être interrogé, lorsque le ménage a été tiré, est alors de 1. Comme tous les ménages ont, par définition, la même probabilité d'être tirés, tous les individus auront eux aussi la même probabilité de faire partie de l'échantillon. Toutefois, si celui-ci est statistiquement satisfaisant, les réponses risquent de ne pas être indépendantes, présentant ce qu'on appelle un *effet de grappe*. On appelle *grappe* un ensemble d'unités tirées simultanément : un ménage constitue une grappe d'individus, un département une grappe de communes, une entreprise une grappe d'établissements ou de salariés, un immeuble une grappe de logements, etc. Or les membres d'une même grappe sont souvent plus semblables entre eux qu'ils ne le sont comparés au reste de la population. Si, par exemple, on étudie des opinions, il y a des chances pour que celles des différents membres soient relativement semblables. Sur le plan des comportements, on pourra observer la même similitude dans les activités de loisir ou les voyages. Les salariés d'une même entreprise se connaissent, travaillent dans les mêmes conditions, sont confrontés à des problèmes communs, et cela s'exprime évidemment dans leurs réponses. Bien que l'importance de ces effets de grappe et de leur retentissement sur l'échantillon total soit, en général, difficile à évaluer avec précision, on peut admettre que le risque existe souvent.

En pratique, on préfère plutôt éviter les effets de grappe, quitte à courir le risque d'un biais lié à la sur-représentation en personnes appartenant à des unités de petite taille, biais considéré, dans la plupart des cas, comme peu important. Le problème est alors de savoir *qui* interroger dans le ménage. On ne peut pas laisser le choix à l'enquêteur, ni aux membres présents du ménage : on risquerait, par exemple, d'interroger systématiquement le chef du ménage, ou son épouse si elle ne travaille pas et est donc plus accessible. La *méthode de Kish* est un procédé aléatoire de désignation de la personne à interroger de manière à éviter de tels biais.

Elle consiste à recenser tous les membres du ménage faisant partie de la population étudiée, et par conséquent susceptibles d'être inclus dans l'échantillon. On en dresse une liste où toutes ces personnes sont ordonnées selon une règle fixe ; par exemple, on énumérera les hommes, puis les femmes, et dans chaque catégorie on ordonnera par âge croissant. Chaque membre du ménage se voit ainsi attribuer un rang.

Pour la désignation de la personne à interroger, l'enquêteur dispose du tableau suivant :

EFFECTIF DU MÉNAGE	RANG DE LA PERSONNE A INTERROGER							
	A	B	C	D	E	F	G	H
1	1	1	1	1	1	1	1	1
2	1	2	2	1	1	2	2	1
3	2	3	3	1	1	3	2	1
4	2	3	4	1	1	4	3	2
5	3	3	5	2	1	5	4	2
6 ou plus	3	5	6	2	1	5	4	2

Les questionnaires ont été attribués d'avance, de façon aléatoire, aux catégories A, B, C, etc. L'enquêteur n'a donc qu'à lire dans le tableau le rang de la personne qu'il doit interroger. Ainsi, si le ménage comprend 4 personnes, et si le questionnaire a été marqué E, il interrogera la personne de rang 1. Ce tableau est conçu de manière que tous les membres du ménage aient approximativement la même probabilité de faire partie de l'échantillon.

L'existence de grappes permet de saisir des populations pour lesquelles il n'existe pas de base de sondage, mais qui font partie de grappes pour lesquelles il en existe une. Par exemple, si l'on veut constituer un échantillon représentatif d'adolescents de 14 à 18 ans, on pourra tirer des listes électorales, considérées comme une approximation acceptable de la population totale, un échantillon d'électeurs qu'on traitera comme un échantillon de ménage. On interrogera alors tous les adolescents qui en font partie, l'échantillon ainsi obtenu étant représentatif.

2.2.1.4. Les echantillons stratifiés

Nous avons vu, en parlant du problème du choix de la population à étudier, qu'un échantillon représentatif de l'ensemble de la population risque de ne comprendre qu'un trop petit nombre de personnes appartenant à certaines catégories, numériquement peu nombreuses, mais importantes pour le problème étudié. Toutefois, ce point n'est en général pas présenté comme un problème de choix de la population à interroger, mais de construction de l'échantillon. En effet, cela revient à constituer l'échantillon en prenant des *taux de sondage* différents selon les catégories considérées (le taux de sondage est la proportion d'individus faisant partie de l'échantillon par rapport à la population, ou à la sous-population, dont ils sont extraits). Un tel échantillon est dit *stratifié*, les catégories en question étant les *strates*. L'échantillon total n'est plus représentatif, mais chacun des sous-échantillons des différentes strates l'est.

Cette manière de faire présente plusieurs avantages. Tout d'abord, elle permet de disposer d'échantillons suffisants pour pouvoir procéder à des analyses approfondies

de chaque strate séparément. C'est particulièrement important lorsqu'on fait l'hypothèse que certains facteurs explicatifs, par exemple, puissent être différents d'une catégorie à l'autre.

Deuxièmement, cela permet une meilleure estimation de certaines grandeurs. Par exemple, si l'on veut connaître le nombre de Français qui prennent des vacances à l'étranger, un échantillon représentatif de l'ensemble de la population comprendra une proportion trop faible de personnes parties à l'étranger pour que la proportion de l'estimation soit satisfaisante. On stratifiera alors en prenant un taux de sondage supérieur dans les catégories dont on suppose qu'elles vont beaucoup à l'étranger : jeunes, cadres, retraités aisés, etc. Certes, en procédant de cette manière, on ne pourra pas prendre immédiatement comme estimation de la grandeur cherchée la valeur de celle-ci dans l'échantillon total, comme on le ferait avec un échantillon représentatif non stratifié. S'il est stratifié comme nous l'avons indiqué, en prenant un taux de sondage supérieur dans les catégories où l'on suppose que le comportement étudié est le plus fréquent, cette valeur globale sera évidemment surestimée. Il est nécessaire de redresser l'échantillon, c'est-à-dire de donner aux observations de chaque strate leur valeur réelle, en pondérant par un coefficient inverse du taux de sondage.

L'avantage de cette procédure est de diminuer l'erreur sur les estimations. En effet, cette erreur, nous le verrons plus loin (cf. 2.2.3.), dépend du nombre d'observations ; il y a donc avantage, pour estimer la fréquence d'un comportement, à le faire à partir du plus grand nombre possible d'observations de ce comportement. Cela peut se faire soit avec un échantillon représentatif non stratifié de grande taille, ce qui est inutilement coûteux, soit en stratifiant de manière à augmenter le nombre de personnes présentant le comportement étudié. Cette deuxième méthode est évidemment plus économique.

Enfin, le dernier avantage de la stratification, c'est de permettre des comparaisons entre groupes. Les tests de comparaison de moyennes, par exemple, ou de proportions, sont les plus sensibles lorsque le nombre d'observations dans chacun des groupes comparés est le même.

2.2.2. Les échantillonnages sur place

Lorsqu'on s'intéresse à une population restreinte, pour laquelle il n'existe pas de base de sondage spécifique, comme, par exemple, les acheteurs d'un certain produit, ceux qui ont voté pour un certain parti, ou une catégorie professionnelle particulière, on peut constituer un échantillon de l'ensemble de la population, par tirage au sort ou par quota, et ne conserver que ceux qui appartiennent à la catégorie visée. Si l'on a pris, en plus, la précaution d'enregistrer le nombre de contacts inutiles, et éventuellement les caractéristiques des personnes non couvertes par l'enquête, on disposera en outre de données comparatives permettant, s'il le faut, de situer la catégorie étudiée par rapport à l'ensemble de la population. Mais cela n'est possible que pour des catégories relativement importantes en volume ; sinon le nombre de contacts nécessaires pour constituer un échantillon de taille suffisante sera trop élevé.

On peut aussi s'appuyer sur le fait que certaines personnes se rencontrent nécessairement dans des endroits particuliers : les automobilistes se trouvent sur les routes ou

dans les stations-service, les voyageurs aériens dans les avions ou dans les aéroports, les fumeurs dans les bureaux de tabac, les clients d'un magasin dans ce magasin, les chômeurs à l'ANPE, etc.

Lorsqu'on peut se contenter d'un échantillon de cette sous-population seule, sans groupe de comparaison, il est possible de constituer un échantillon correct en se rendant dans de tels lieux et en procédant, sur place, à un tirage au sort parmi les personnes présentes. Par exemple, pour obtenir un échantillon d'automobilistes, on pourra se poster près d'une pompe à essence, et interroger les conducteurs qui s'y arrêtent. Selon l'échantillon qu'on veut obtenir, on les abordera tous, ou seulement un sur deux, ou un sur cinq, etc.

La méthode est en principe satisfaisante, mais elle n'est pas sans biais ; l'exemple de la constitution d'un échantillon d'automobilistes par contact aux pompes à essence nous permettra de les recenser.

En premier lieu, il est évident qu'on ne peut pas se contenter de se placer auprès d'une seule pompe : ce ne sont certainement pas les mêmes conducteurs qu'on rencontrera dans différents quartiers, en ville, ou dans un village, ou en rase campagne, sur une petite route ou sur une autoroute. Il faudra donc commencer par constituer un échantillon représentatif des pompes à essence de la région étudiée, ou au moins constituer un échantillon raisonné, comprenant une variété suffisante de cas.

A ce premier *échantillonnage spatial*, il faut ajouter un *échantillonnage temporel*. En effet, les automobilistes qui se déplacent de jour sont différents, de bien des points de vue, de ceux qui se déplacent de nuit, ceux qui circulent en semaine différents de ceux qui circulent le samedi ou le dimanche. Il faudra donc soit effectuer l'enquête pendant toute une semaine, soit choisir, par tirage au sort ou autrement, des moments divers.

L'importance de ces précautions dépendra de la nature du problème étudié. Si l'enquête porte, par exemple, sur la sécurité, on peut penser, effectivement, que les conducteurs de jour ou de nuit auront des réactions différentes. Répartir les enquêtes sur vingt-quatre heures sera donc nécessaire. En revanche, si l'objet de l'enquête est l'attitude à l'égard des différentes marques d'essence, cette précaution sera superflue : on n'a guère de raison de penser que les préférences des conducteurs de jour sont très différentes de celles des conducteurs de nuit. Mais, dans ce dernier cas, il faudra accorder une attention particulière au choix des pompes retenues.

L'échantillonnage spatial et temporel permet d'éliminer un certain nombre de biais, mais il n'assure pas nécessairement un échantillon représentatif. Rappelons-en la définition : il s'agit d'un échantillon constitué de manière telle que tous les membres de la population aient la même probabilité d'en faire partie. Dans la plupart des cas d'échantillonnage sur place, cette condition n'est pas remplie : tous les membres de la population ne passent pas nécessairement entre ces points avec la même fréquence, ce qui introduit des distorsions, puisque *la probabilité de faire partie de l'échantillon est proportionnelle à la fréquence de passage*.

Dans le cas des automobilistes, la probabilité de s'arrêter à une pompe à essence est d'autant plus élevée qu'on circule plus. L'échantillon ainsi constitué sera donc surreprésenté en personnes se déplaçant beaucoup. D'une façon générale, tout ce qui modifie la consommation d'essence biaise en principe l'échantillon : la vitesse à laquelle

on circule, la consommation au kilomètre, la capacité du réservoir, etc. Certains de ces biais peuvent être faibles ou s'annuler approximativement, mais pas tous, et il n'est pas possible de les éliminer dans la constitution de l'échantillon ; en revanche, on peut recueillir les informations nécessaires pour effectuer un redressement. Par exemple, si l'on estime que le biais le plus important est causé par les différences de kilométrage parcouru, on pourra demander à chaque conducteur interrogé combien il fait en moyenne de kilomètres par semaine, ou par mois, et redresser l'échantillon en accordant à chacun un poids inverse du nombre de kilomètres parcourus.

Ces distorsions liées à la fréquence ne se produisent pas toujours. On pourra constituer, par exemple, un échantillon représentatif de chômeurs en prenant contact avec eux dans les bureaux où ils vont « pointer ». Comme leur passage y est obligatoire la seule précaution à prendre porte sur le choix des bureaux, en tenant compte du fait que les cadres pointent toutes les deux semaines, et les non-cadres toutes les semaines.

Si l'échantillonnage sur place permet en principe d'obtenir, moyennant les précautions dont nous avons parlé, des échantillons acceptables, la méthode se heurte à des difficultés pratiques importantes. Tout d'abord, la plupart des endroits en question ne sont pas des lieux publics ; il faut donc obtenir l'autorisation de s'y poster, autorisation qui peut être refusée pour diverses raisons : crainte d'effrayer ou d'importuner une clientèle, méfiance à l'égard de l'enquêteur ou de l'objectif de l'enquête, etc.

De plus, les personnes ainsi abordées sont rarement disponibles plus de quelques minutes ; en outre, le cadre se prête mal, en général, à la passation d'un questionnaire, et encore moins à des entretiens : on est le plus souvent debout, l'environnement est bruyant, etc. C'est pourquoi on utilise uniquement des questionnaires très courts, de l'ordre de cinq à sept minutes au maximum, ce qui fait au plus une dizaine de questions simples. Si on tient à utiliser un questionnaire plus approfondi, on peut aussi aborder les personnes en leur demandant d'accepter de recevoir un enquêteur à leur domicile, bien que cela éveille parfois la méfiance de gens qui répugnent à donner leur adresse à des inconnus.

2.2.3. Les échantillonnages par quota

Devant l'impossibilité, dans la majorité des cas, de disposer d'une base de sondage satisfaisante et adaptée au problème traité, et devant les difficultés pratiques de la mise en œuvre de méthodes d'échantillonnage spatial, on se trouve souvent contraint à renoncer à des méthodes aléatoires rigoureuses ; on peut alors contourner la difficulté par le recours à la méthode dite *d'échantillonnage par quota*, qui est de loin, actuellement, la plus utilisée.

Dans son principe, la méthode des quota consiste à obtenir une représentativité suffisante en cherchant à reproduire, dans l'échantillon, les distributions de certaines variables importantes, telles que ces ditributions existent dans la population à étudier. Par exemple, si celle-ci comporte autant d'hommes que de femmes, on fera en sorte qu'il en soit de même dans l'échantillon ; si la population en question comporte 40 % d'ouvriers, on s'efforcera d'avoir 40 % d'ouvriers dans l'échantillon interrogé, etc.

Ces « quota » sont imposés aux enquêteurs qui sont, à part cette contrainte et éven-

tuellement quelques autres portant sur les lieux et les moments des enquêtes, et sur lesquelles nous reviendrons, libres du choix des personnes à interroger.

En général, les quota sont définis en fonction de quelques caractéristiques simples, comme l'âge, le sexe, la catégorie socio-professionnelle, etc. Ces variables sont retenues en premier lieu parce qu'on peut en connaître la distribution dans la population ; il est évident que c'est une condition nécessaire pour l'application de la méthode. Ensuite, il faut que ce soient des caractéristiques faciles à connaître : s'il faut trop de temps, trop de questions préalables, pour déterminer à quelle catégorie un sujet potentiel appartient, les contacts inutiles risquent d'être trop longs, augmentant ainsi la durée et le coût de l'enquête. Il faut aussi que les variables retenues soient effectivement en corrélation, soit directement avec les comportements qu'on veut étudier, soit avec d'autres facteurs qui interviennent dans les hypothèses. Sauf cas particuliers, ce sont les variables dites « classiques », âge, sexe, catégorie socio-professionnelle, qui remplissent toutes ces conditions.

Enfin, il faut que les quota soient fixés de manière à corriger les distorsions dues au fait que certaines catégories sont plus accessibles que d'autres. Par exemple, les femmes qui travaillent à l'extérieur sont moins souvent chez elles que celles qui restent au foyer. Si l'on fait du porte-à-porte, elles seront plus difficiles à rencontrer. Si l'on fixe simplement le nombre de femmes à interroger, l'échantillon sera biaisé, car il contiendra une trop faible proportion d'actives. On fixera donc des quota distincts pour les deux catégories de femmes.

Lorsqu'on fixe les quota, on peut soit traiter les différentes variables de façon indépendante, soit les « croiser ». Par exemple, si les quota portent sur l'âge et le sexe, on pourra fixer indépendamment les proportions d'hommes et de femmes d'une part, et celles des différentes classes d'âge d'autre part, en espérant que la répartition se fera à peu près correctement, c'est-à-dire, par exemple, qu'on n'aura pas une majorité de femmes chez les jeunes et une majorité d'hommes chez les plus âgés. Si l'on craint que de telles distorsions puissent se produire, on « croisera » les critères, c'est-à-dire qu'on imposera les proportions d'hommes et de femmes dans chaque classe d'âge.

Il est inutile d'accumuler les contraintes, que ce soit par le nombre des critères retenus, ou par leur croisement, pour s'assurer d'un « meilleur » échantillon. Tout d'abord, cela complique exagérément le travail des enquêteurs. Rechercher un employé, célibataire, de 36 - 45 ans, habitant en pavillon, devient une tâche presque impossible, en tout cas inutilement longue. Tous ceux qui ont une certaine expérience de la technique connaissent ces « queues d'enquêtes » où l'on attend avec impatience les quelques derniers cas, à peu près introuvables. Il est préférable de s'en tenir à un nombre de critères limités, de ne les croiser que si on a des raisons sérieuses de craindre une distorsion, en sachant que si les quota sur les critères importants sont respectés, il y a de fortes chances pour que l'échantillon soit à peu près représentatif selon les autres variables. Il est en revanche plus important de se prémunir contre les biais liés aux refus et aux absences, dont nous parlerons en détail au chapitre suivant.

La qualité d'un échantillon par quota dépend de la manière dont les enquêteurs recherchent les sujets et prennent contact avec eux. Pour assurer une meilleure représentativité, on associe la méthode des quota avec une méthode aléatoire, comme le *son-*

dage aérolaire, par exemple. Cela consiste à tirer au sort certaines zones dans l'agglomération, et à envoyer les enquêteurs dans ces zones. Là, ils pourront soit accoster les passants, soit faire du porte-à-porte, soit éventuellement combiner les deux. Si l'on recherche des mères de famille, on pourra aller dans les squares ou à la sortie des écoles, mais en sachant que, de cette manière, on ne saisit que des mères qui ne travaillent pas à l'extérieur, et qui ont des enfants très jeunes.

Le porte-à-porte est la meilleure méthode. Ignorant qui va ouvrir la porte à laquelle il a sonné, l'enquêteur n'est pas influencé par l'air plus ou moins aimable, auquel il sera certainement sensible s'il choisit ses sujets dans la rue. D'autre part, la passation du questionnaire à domicile est indispensable dès qu'il dépasse cinq à sept minutes. En venant dans la journée, on risque de ne rencontrer que des femmes, des personnes âgées, éventuellement quelques jeunes. Ce sont donc de très bonnes heures pour rencontrer ces catégories, mais si l'on veut aussi interroger des femmes ou des hommes actifs, il faudra également venir le soir ou pendant le week-end.

La méthode des quota présente l'avantage d'être toujours applicable, quelle que soit la population étudiée, et que l'échantillon qu'on vise soit simplement représentatif, stratifié, ou conçu selon un plan expérimental plus ou moins complexe. Son coût est moins élevé que celui d'une enquête sur adresses [1]. En revanche, elle ne donne pas des échantillons aléatoires, au sens statistique du terme. Il en résulte, en particulier, qu'on n'est pas dans la situation postulée par les différentes techniques d'inférence statistique (tests d'hypothèses, estimations, calculs d'erreur, etc.). En effet, en imposant les quota, on introduit une contrainte forte, dont les techniques statistiques ne tiennent pas compte, et d'autre part on sait mal construire un modèle probabiliste du procédé de choix des individus par les enquêteurs.

La qualité d'un échantillon par quota dépend plus des procédés effectivement employés par les enquêteurs pour choisir leurs sujets que de la rigueur des consignes qui leur sont données. Il est inutile d'imposer des quota très stricts si les enquêteurs ne prennent pas garde aux problèmes d'échantillonnage spatial ou temporel, ou si la proportion de personnes qui refusent est très forte.

La reproduction des distributions de la population doit être considérée comme une condition nécessaire, mais non suffisante, de la qualité d'un échantillon. Le respect des quota ne garantit pas l'absence de biais systématiques.

Il y a quelques années (nous ignorons si c'est encore le cas aujourd'hui) les caractéristiques des lecteurs de *Paris-Match* (âge, sexe, catégorie socio-professionnelle) étaient très semblables à celles de l'ensemble des Français. Un échantillon de ces lecteurs aurait donc constitué un bon échantillon par quota de la population française. C'est pourtant une conclusion difficile à accepter : on a l'impression que les lecteurs de *Paris-Match* doivent se distinguer des non-lecteurs sur d'autres points importants.

Malgré les apparences, un échantillon par quota est donc beaucoup moins maîtrisé qu'un échantillon aléatoire. Ceux-ci, à première vue, laissent plus de marge aux erreurs d'échantillonnage, mais celles-ci sont calculables, et on sait comment en tenir compte. Un échantillon par quota ne présente, en apparence, aucune erreur sur les variables

1. On admet qu'un enquêteur, dans un temps donné, fait environ 50 % d'enquêtes en plus lorsqu'il travaille par quota plutôt que sur adresses.

contrôlées, mais ce ne sont pas les seules connaissances dont on dispose. Théoriquement, on ne se trouve pas dans un cas où l'on peut utiliser les techniques statistiques. En fait, on les traite très souvent comme des échantillons vraiment aléatoires, et il ne semble pas que les conclusions qu'on en ait tirées soient mauvaises. Toutefois, elles peuvent parfois être suspectes, sans qu'on sache précisément comment traiter le problème.

2.2.4. Les biais dans la réalisation de l'échantillon

2.2.4.1. Vue d'ensemble

La qualité des conclusions qu'on peut tirer d'une enquête est liée à la composition de son échantillon. Si celui-ci n'est pas représentatif de la population particulière qu'on veut étudier, les conclusions qu'on en tirera, même au moyen de méthodes statistiques raffinées, s'écarteront de celles qu'on aurait obtenu en interrogeant toute la population. L'échantillon est biaisé, la généralisation n'est pas légitime.

Ces biais ont différentes origines. Tout d'abord, certains peuvent s'introduire au moment même de la constitution de l'échantillon. Par exemple, faute d'une liste exhaustive, on prend, comme approximation, une base de sondage qui ne couvre qu'une partie de la population visée. C'est le cas lorsqu'on utilise les listes électorales pour étudier l'ensemble de la population. Ou, lorsqu'on cherche à étudier la population qui fréquente un équipement quelconque en effectuant un tirage au sort parmi ceux qui y vont, sachant qu'une personne a d'autant plus de chances de faire partie de l'échantillon qu'elle y va plus souvent. Nous avons vu ces différents problèmes à propos de la constitution des échantillons.

Mais se garder de ces biais, ou « redresser » l'échantillon si l'on n'a pas pu les éviter, ne constitue pas une garantie suffisante. Disposer d'une base de sondage correcte, ou avoir tiré une liste de noms par un procédé vraiment aléatoire, nous donne sur le papier un échantillon parfait. Il reste cependant à prendre contact avec les personnes ainsi désignées par le sort, et à les interroger. Or, il n'y a pas nécessairement coïncidence entre la liste initiale et celle des personnes effectivement interrogées ; et ceci pour deux raisons : certaines personnes ne sont pas accessibles pour des raisons diverses (absence temporaire, voyage, maladie, déménagement, etc.) et d'autres refusent de se prêter à l'enquête. Malgré quelques différences, que nous préciserons, les mêmes problèmes se posent pour les échantillons par quota.

En principe, tout écart entre les personnes qui auraient dû être interrogées, et celles qui le sont effectivement, peut être source de biais. En fait, celui-ci existera chaque fois que *les personnes non interrogées auraient donné des réponses différentes de celles qui l'ont été.* Par définition, il est impossible de s'en assurer directement. Néanmoins, en connaissant certaines des caractéristiques des personnes inaccessibles ou ayant refusé de répondre, on peut avancer quelques hypothèses sur ces écarts systématiques.

Nous examinerons en détail les problèmes posés par les personnes absentes et par les refus dans le cas des échantillons aléatoires, puis nous passerons au cas des échantillons par quota, où l'on rencontre les mêmes difficultés, et quelques autres en plus.

2.2.4.2. Les biais d'échantillonnage dans les échantillons aléatoires

a. *Les absences*

Une personne peut se trouver absente de son domicile lors du passage d'un enquêteur pour de multiples raisons : elle est sortie pour faire des courses, elle travaille, elle est hospitalisée, elle est en voyage d'affaires ou en vacances, etc. Si l'enquêteur passe immédiatement à la personne suivante sur la liste, en se disant qu'après tout elle a été, elle aussi, tirée au sort, on obtiendra un échantillon de personnes sortant peu. Plus précisément, on aura d'autant moins de chances de faire partie de l'échantillon final qu'on sort plus. Cela implique, par exemple, qu'on trouvera plus de femmes ne travaillant pas à l'extérieur que de femmes actives ou d'hommes, plus de personnes âgées que de jeunes, plus de personnes n'ayant pas de voiture que de personnes motorisées, etc. Comme tous ces groupes ont vraisemblablement sur beaucoup d'autres points des comportements et des attitudes qui diffèrent, les conclusions qu'on pourra tirer d'un tel échantillon seront certainement biaisées.

Pour éviter ces distorsions, certaines précautions sont habituelles. Par exemple, on évite d'effectuer des enquêtes pendant les mois de juillet et d'août, à cause des vacances, les personnes qui partent étant certainement différentes de celles qui restent.

En outre, un enquêteur à qui on a donné le nom et l'adresse d'une personne à interroger ne devrait jamais renoncer après avoir simplement constaté son absence ; il est indispensable qu'il revienne à une autre heure, ou un autre jour. En général, on impose aux enquêteurs de n'abandonner une adresse qu'après trois, ou même quatre, visites infructueuses, et après avoir tenté de se renseigner auprès d'autres personnes du ménage, de la concierge, éventuellement des voisins, du moment où il a le plus de chances de rencontrer la personne en question, ce qui peut l'amener à revenir tard le soir, ou le week-end.

Il peut aussi laisser une formule avec une enveloppe-réponse, en demandant à l'intéressé de fixer lui-même un rendez-vous. Toutefois, il ne faut pas surestimer cette méthode : le taux de réponse peut être faible ; une personne qui accepterait de recevoir un enquêteur qui se présente peut négliger de renvoyer un papier.

On peut aussi essayer de prendre un premier contact par téléphone, ce qui évite des déplacements éventuellement longs. Mais il faut être conscient du fait que, ce faisant, l'échantillon risque de comprendre une trop forte proportion de personnes ayant le téléphone dont la possession est liée à la catégorie sociale.

Malgré toutes ces précautions, il y a des personnes qu'on ne trouve pas. Leur proportion n'est d'ailleurs pas très élevée : entre 1 % et 5 % sur un échantillon de la population urbaine française, lorsqu'on prend toutes les précautions que nous avons mentionnées. Doit-on craindre que ce petit groupe soit différent de ceux avec lesquels on a pu prendre contact ? Ils se caractérisent vraisemblablement, en moyenne, par de plus longues absences du domicile, au moins au moment de l'enquête. On trouvera parmi eux, par exemple, des célibataires qui sont rarement chez eux. Cela implique certainement un mode de vie particulier, donc probablement des comportements, des attitudes, des opinions spécifiques. Les personnes actives, travaillant hors du domicile, sont moins faciles à atteindre que celles qui ne travaillent pas ou qui travaillent à domicile.

Les jeunes sortent plus souvent le soir, ou pendant le week-end, que les plus âgés, les cadres plus que les ouvriers. Une personne que son travail amène à se déplacer souvent sera, elle aussi, peu accessible. Tout ceci donne des indications sur la nature du biais introduit par ces absences. Comme elles sont de natures diverses et, dans l'ensemble, numériquement faibles, on peut admettre qu'elles ne mettent pas radicalement en cause la valeur d'une enquête.

On a proposé une méthode ingénieuse de correction de l'échantillon pour prendre en considération les personnes absentes. Elle consiste à comparer les réponses données par les personnes qu'on a pu joindre lors de la première tentative, à celles des personnes qu'on n'a pu trouver qu'après deux ou trois essais. On peut ainsi construire la courbe de la proportion d'une certaine réponse en fonction du nombre de tentatives nécessaires pour rencontrer la personne. Par extrapolation, on en déduit ce qu'auraient répondu les personnes inaccessibles (cf. fig. 1). Si ces différents groupes donnent les mêmes réponses, on pourra en conclure que les absences n'introduisent pas de biais dans l'échantillon.

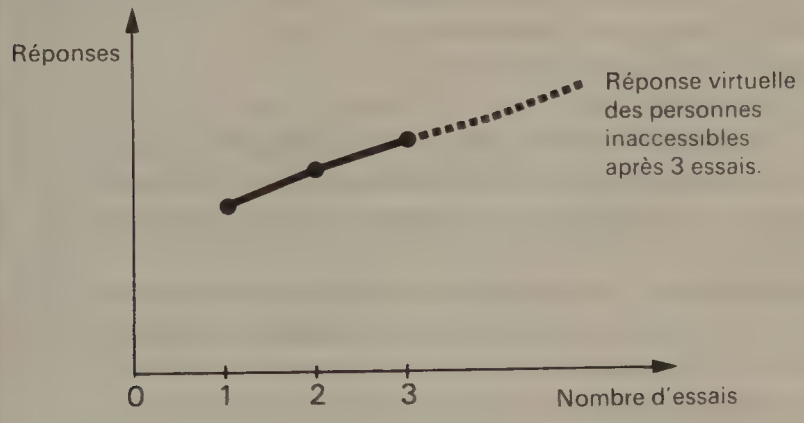


Fig. 1

A côté de ces absences temporaires, permettant en principe à l'enquêteur de rencontrer plus tard la personne désignée, il y a des absences de longue durée, qui se prolongent au-delà de la durée de l'enquête (voyages, hospitalisation), et aussi les absences définitives, les déménagements en particulier, ou les décès. En effet, les bases de sondage dont on peut disposer, par exemple les listes électorales, ne sont pas nécessairement à jour. En France, rien n'oblige à déclarer un changement de domicile, et certaines personnes continuent à être inscrites à la Mairie d'un domicile qu'elles ont quitté depuis plusieurs années. Retrouver ces personnes est très souvent difficile. Non seulement on a de la peine à connaître leur nouvelle adresse, mais celle-ci peut être située en dehors de la zone d'enquête. Dans ce cas, on considère comme tout à fait acceptable de remplacer la personne désignée par le tirage au sort par une autre, habitant le même logement. Il y a en effet des chances pour que des ménages qui se succèdent dans un même logement présentent de nombreuses caractéristiques communes.

Il en va de même pour les décès. Ceux-ci sont en général assez rapidement enregistrés sur les listes électorales, mais peut-être pas sur d'autres listes. Là encore, le remplacement par une personne habitant le même logement se justifie.

La rénovation urbaine, ou la construction de quartiers nouveaux, peuvent être sources de biais. On ne trouve souvent sur les bases de sondage disponibles que les anciennes adresses, alors que celles-ci n'existent plus. Il n'y a donc pas d'occupants du même logement, et ce peut être une catégorie importante de la population qui devient inaccessible. Il n'est pas possible d'indiquer de règle générale permettant de réduire de tels biais. C'est dans chaque cas particulier qu'il faut trouver la solution : rechercher d'autres listes, faire des enquêtes complémentaires par quota, etc.

Devant toutes les formes d'absences, qu'elles soient temporaires ou définitives, il reste toujours une possibilité, qu'on peut appliquer dans tous les cas : remplacer la personne qu'on n'a pas pu atteindre par une autre, également tirée au sort, et appartenant à la même catégorie (âge, sexe, catégorie socio-professionnelle). Ceci suppose en premier lieu qu'on puisse connaître ces caractéristiques avant de rencontrer la personne en question, donc qu'elles soient indiquées sur la liste qui a servi au tirage. En outre, on apporte à l'échantillonnage aléatoire quelques traits de la méthode des quota, ce qui y introduit les problèmes propres à cette méthode. Néanmoins, comme les différentes catégories qui composent la population sont très inégalement accessibles, cette solution est préférable au choix arbitraire d'une autre personne.

b. *Les refus*

Sauf pour le recensement, et pour quelques enquêtes de l'INSEE, nul, en France, ne peut être contraint à se prêter à un entretien, ni à répondre à un questionnaire. Effectivement, dans pratiquement toutes les enquêtes, on constate qu'un certain nombre de personnes refusent de répondre.

Les refus sont le plus souvent globaux : la personne contactée refuse de recevoir l'enquêteur. En revanche, il est très rare que quelqu'un ayant accepté de le recevoir arrête au milieu du questionnaire.

On dispose de relativement peu de données sur l'importance de ces refus : rares sont les comptes rendus d'enquête où l'on indique le nombre (et les caractéristiques) des personnes qu'il a fallu contacter pour atteindre un échantillon de la taille voulue. Cela est très regrettable, car, comme nous le verrons, on a tout lieu de supposer que les refus introduisent des biais non négligeables. Pour apprécier ceux-ci il faut répondre, au moins de façon hypothétique, à la question : les personnes qui ont refusé l'enquête auraient-elles répondu de façon différente de celles qui ont accepté ?

Pour apporter des éléments de réponse, il faut distinguer différents types de refus. En premier lieu, on peut distinguer ceux qui s'opposent à tout type d'enquête de ceux qui refusent à cause du thème lui-même. Les biais risquent évidemment d'être plus importants dans le second cas, bien qu'ils ne soient pas négligeables dans le premier.

Mentionnons quelques formes de refus global, celui qui s'exprime avant même de connaître le thème de l'enquête.

1. *Les travailleurs indépendants, commerçants, artisans*, refusent plus souvent que d'autres catégories sociales. Ils sont souvent très occupés ou, ayant des horaires mal définis, ne se rendent jamais disponibles.

D'autre part, l'idéologie propre à cette catégorie sociale les amène souvent à faire une distinction très nette entre le « public » et le « privé » et à considérer toute ques-

tion comme une intrusion dans le « privé ».

2. *Les personnes âgées* : elles sont craintives, hésitant à laisser un étranger pénétrer chez elles. D'autre part, elles ont très souvent une attitude de retrait, de désengagement par rapport au monde, attitude qui les amène à refuser leur participation.

3. *Certaines femmes mariées*, qui refusent en se retranchant derrière leur mari, en disant soit qu'elles ne peuvent pas accepter sans son accord, soit que lui-même saura mieux répondre.

4. *Certains intellectuels* refusent systématiquement par hostilité au principe même des sondages d'opinion, auxquels ils assimilent d'ailleurs toutes les formes d'enquêtes. Ce type de refus semble croître actuellement.

5. *Certains milieux politisés* s'opposent également de façon systématique aux enquêtes ; la nature de cette opposition paraît d'ailleurs varier au cours du temps.

Enfin, on constate que le taux de refus croît très nettement avec la taille de la ville : il est maximum dans la région parisienne.

A part ces refus qu'il est possible de localiser dans certaines catégories sociales, les autres raisons invoquées sont le manque de temps, la crainte de laisser pénétrer un inconnu chez soi. Il semblerait aussi que chez des personnes de bas niveau socio-culturel la crainte de « ne pas être à la hauteur », de ne pas « savoir répondre », soit une cause de refus. Mentionnons également la crainte, justifiée, que l'enquêteur ne soit en réalité un représentant qui utilise ce prétexte pour s'introduire ; cette pratique est en effet fréquente, et cause certainement un grand tort aux véritables enquêteurs.

Malgré tous ces refus, la majorité de la population accueille bien les enquêteurs. Les débutants sont même souvent surpris de constater la satisfaction de certaines personnes, heureuses qu'on leur demande leur avis, qu'on les écoute. Rares sont ceux qui ont souvent la possibilité de s'exprimer, et l'enquête leur en fournit une occasion. De plus, les enquêtes qui, encore une fois, sont en général assimilées aux seuls sondages d'opinion, jouissent, dans la majorité de la population, d'un grand prestige : on espère souvent que les résultats seront publiés, et, répondre, c'est une manière d'« être dans le journal ».

On constate toutefois depuis quelques années, aux Etats-Unis, une certaine saturation : les enquêtes sont devenues trop nombreuses, elles ont perdu leur prestige, et les personnes qui, par leur profession par exemple, sont amenées à être sollicitées plusieurs fois, s'en lassent et on observe que le taux global de refus augmente.

En France, on n'en est pas encore là, mais on peut s'attendre à une évolution identique au cours des prochaines années.

Devant des refus en proportion « raisonnable », c'est-à-dire ne dépassant pas 15 à 20 %, on réagit en général en remplaçant une personne ayant refusé par une autre présentant les mêmes caractéristiques, comme on le fait pour les absents. Sur un échantillon de l'ensemble de la population, où les causes de refus sont diverses, chacune relativement peu fréquente, on peut admettre qu'on n'introduit pas de cette manière de biais grave. C'est d'ailleurs le principe même de la méthode des quota.

En revanche, lorsqu'on s'adresse à une catégorie particulière, il est vraisemblable que les causes de refus seront moins variées, et par conséquent que les « non-répondants » seront systématiquement différents des « répondants », et en général de

la même façon. Par exemple, au cours d'une enquête auprès de personnes âgées, on constate environ 50 % de refus. Les raisons invoquées (« ça ne m'intéresse pas », « on ne fait rien pour moi, je ne ferai rien pour vous », « je n'ai rien à dire à personne ») suggèrent que ces refus sont liés à une des variables de base qui distingue ces personnes : le degré de désengagement. Les personnes les plus « désengagées », celles qui sont, d'une certaine façon, coupées des relations avec le monde, ont plus de chances de refuser que les autres. Dans la mesure où cette variable est certainement un des facteurs explicatifs les plus discriminants du comportement des personnes âgées, le biais ainsi introduit est évident, et ses dangers sont certains.

Pour réduire le nombre des refus plusieurs méthodes sont utilisées. Une lettre d'introduction annonçant le passage de l'enquêteur et précisant le but de l'enquête, envoyée quelques jours à l'avance, augmente le taux d'acceptation. L'enquêteur n'est plus un inconnu qui vient frapper à la porte de façon inopinée, c'est le représentant d'une institution. L'influence de cette lettre dépend évidemment de l'organisme qui l'envoie. L'en-tête d'un bureau d'études inconnu n'aura qu'un effet modéré. Un institut de sondage connu, dont on voit le nom dans les journaux, lève plus de réticences. La recommandation d'organismes plus connus, surtout publics, réduisent considérablement le nombre de refus. Nous avons pu le constater lors d'une enquête auprès d'un échantillon de personnes âgées. Avec une lettre d'introduction d'un bureau d'études n'ayant aucune notoriété, le taux de refus était proche de 60 %. Une lettre de la mairie de la commune l'a fait tomber à 4 % et les seules personnes ayant refusé étaient de grands malades, physiquement incapables de recevoir un enquêteur.

Certains refus sont provoqués par des circonstances momentanées, ou par un enquêteur qui inspire de la méfiance, ou qui est antipathique à certaines personnes. On peut réduire ce type de refus en envoyant un autre enquêteur, d'aspect différent, quelques jours plus tard. C'est une méthode lourde, mais qui semble efficace : le nombre de refus peut être diminué de moitié environ.

2.2.4.3. Les biais d'echantillonnage dans les enquêtes par quota

Lorsqu'on constitue un échantillon statistique par tirage au sort à partir d'une base de sondage, on obtient une liste bien définie d'adresses de personnes à interroger. Si l'une d'elles est inaccessible, ou si elle refuse, on peut le noter, et l'écart entre l'échantillon initial et l'échantillon réel peut être connu avec précision, ce qui permet d'estimer son incidence sur la validité des conclusions qu'on peut en tirer.

Dans les enquêtes par quota, la situation est beaucoup moins claire. Les enquêteurs ne disposent pas d'une liste de noms, mais uniquement de consignes quant à la composition de l'échantillon. Quelles que soient les difficultés qu'ils rencontrent, ils finiront par rapporter un ensemble de questionnaires remplis présentant les caractéristiques voulues, où ne manqueront, au pire, que quelques personnes appartenant à des catégories très peu nombreuses.

On peut donc avoir facilement l'illusion que tout va bien : puisque l'échantillon interrogé présente bien la composition voulue, il peut être considéré comme représentatif. Pourtant, en réalité, les risques de biais sont plus importants que dans le cas des échantillons aléatoires, mais ils sont moins visibles.

Les distorsions de l'échantillon dues aux absences ou aux refus existent, quelle que soit la méthode utilisée. Mais elles sont plus difficiles à saisir lorsqu'on utilise la méthode des quota. En effet, un enquêteur qui fait du porte-à-porte ignore qui se trouve dans chaque logement. S'il n'y a personne, il ne peut pas savoir si celui ou celle qui y habite présente les caractéristiques voulues ; il ne fera donc pas d'effort particulier pour revenir à un autre moment et préférera tenter sa chance à la porte voisine. On ne pourra donc pas savoir s'il s'agit d'une absence, à enregistrer comme telle, de quelqu'un faisant partie d'une catégorie déjà complète, ou d'une personne ne faisant pas partie de la population visée. On rencontre les mêmes problèmes pour les refus, également difficiles à localiser, et qui ne seront pas dénombrés comme tels. De plus, comme, par définition de la méthode, on ne connaît à l'avance ni le nom, ni l'adresse, des personnes qu'on va chercher à interroger, il est difficile de s'annoncer par une lettre pour essayer d'obtenir un meilleur accueil, ce qui peut augmenter le nombre des refus. Les deux sources de biais que nous avons distinguées par les enquêtes statistiques se retrouvent donc dans la méthode des quota, mais il est plus difficile d'en évaluer l'importance.

De surcroît, à ces biais généraux, communs à toutes les formes de sondage, viennent s'ajouter ceux qui sont dus à la manière dont les enquêteurs choisissent les personnes qu'ils vont aborder. Nous avons vu que ces biais peuvent être réduits en donnant des consignes très strictes concernant le lieu et l'heure des enquêtes. Mais cela ne suffit pas ; dans la rue, par exemple, on hésitera à aborder quelqu'un visiblement pressé, ou lourdement chargé. Souvent cela n'a qu'une importance minime, mais il est important de se poser le problème dans chaque cas et de se demander si le biais introduit est vraiment gênant.

Un autre risque de la méthode des quota est de ne saisir, en fait, que des sous-catégories particulières des catégories fixées. Par exemple, on définira des catégories d'âge, avec un groupe de « plus de 65 ans ». Il y a des chances pour que, en fait, on n'obtienne qu'un échantillon de personnes de 65 à 70 ans, les âges supérieurs étant absents ou non représentés, pour des raisons de santé, de refus, etc. Ou encore, cherchant à étudier des ouvriers, on les cherchera à un endroit tel qu'on n'interrogera que des salariés de grosses industries.

Le risque que les enquêteurs n'interrogent que des personnes qu'ils connaissent déjà est grand. Si on ne leur donne pas des consignes précises à ce propos, il est normal qu'ils préfèrent interroger des personnes de leurs relations, plutôt que d'aller faire du porte-à-porte et aborder des inconnus. La distorsion de l'échantillon ainsi obtenu est très difficile à estimer. De plus, lorsque l'enquêteur et l'enquêté se connaissent, la relation qui s'établit entre eux n'est pas la même que si ce sont des inconnus, et cela peut introduire un biais supplémentaire, qui n'est pas lié à l'échantillonnage mais aux réponses données. Nous reviendrons sur ce point dans le chapitre consacré à la relation enquêteur-enquêté et à l'analyse psychologique de la situation d'enquête (cf. 5.3.).

2.2.4.4. L'importance des biais de réalisation de l'echantillon

Passer en revue ces différentes sources de biais, auxquelles il faut ajouter celles qui sont liées aux modes de constitution de l'échantillon, nous a montré qu'il n'existe aucune

méthode qui nous assure dans tous les cas un échantillon absolument représentatif. Toutefois, s'il faut être bien conscient du problème, rester attentif à ces difficultés, et ne pas se satisfaire trop facilement d'un tirage au sort ou du respect du quota, il faut aussi éviter un purisme exagéré et condamner tout sondage à cause de ses biais inévitables.

Tout d'abord, il faut évaluer l'importance de ces biais. Quelques pour-cent de refus ou d'absents n'affectent pas sérieusement la validité des résultats, surtout si c'est la somme de biais divers, qui ont des chances de s'annuler. Les absences de jeunes célibataires et celles de cadres en voyage, les refus de personnes âgées ou de commerçants, sont suffisamment divers pour qu'on puisse admettre que, dans l'ensemble, ceux qui n'ont pas pu être touchés auraient répondu de façon semblable aux autres.

Le problème se pose en revanche de façon aiguë lorsqu'on veut saisir une sous-population particulière. Si, par exemple, on s'intéresse aux mères de famille, les refus ou les absences risquent d'être liés au fait de travailler à l'extérieur, et sur beaucoup de points on peut craindre que les personnes actives ne réagissent de façon différente des autres. Dans ce cas, même si numériquement la population de personnes non touchées est faible, le fait qu'il y ait une cause principale de distorsion risque d'introduire un biais important.

L'importance du biais dépend aussi du problème traité. Supposons qu'on fasse une enquête le soir. Certaines personnes refuseront de répondre, pour ne pas manquer une émission de télévision. Si l'enquête porte précisément sur la télévision, ou sur les loisirs, le biais sera certainement sensible.

Ce sera encore le cas, bien que de façon moins sensible, pour une enquête politique. Mais s'il s'agit, par exemple, de consommation alimentaire, rien a priori ne permet de supposer un lien, même indirect, entre l'attachement à la télévision et la préférence pour tel produit. Même si la proportion de refus est élevée, l'échantillon pourra être considéré comme satisfaisant.

Enfin, la gravité du biais est aussi fonction de la nature des résultats recherchés. Si ce qu'on cherche ce sont des *estimations* de grandeurs (dépenses, consommations diverses, fréquentation, intentions de vote, etc.), tous les biais dont nous avons parlé risquent de fausser les conclusions, et il faudra prendre un soin tout particulier pour les éliminer. C'est le cas pour la plupart des enquêtes réalisées par des économistes. En revanche les sociologues et psychosociologues cherchent souvent plus à mettre en évidence des *relations*, que ce soit sous la forme de corrélations ou sous la forme de comparaisons entre groupes. Le problème est alors très différent. Dans certains cas, un biais dans les distributions de deux variables n'affecte pas leur corrélation. Ou encore, lorsqu'on compare deux groupes, le fait qu'ils ne soient pas représentés dans l'échantillon total avec leur véritable proportion ne change rien à la valeur de la comparaison. En revanche, si un des groupes est biaisé du point de vue d'une troisième variable, ou si les deux le sont inégalement, la comparaison ne sera pas valide.

Prenons comme exemple une enquête sur les intentions de vote, et supposons que l'âge soit une variable importante, dont on veut étudier l'influence. On peut, pour cela, comparer les réponses données par des personnes de différentes classes d'âge, ou, ce qui revient au même, étudier la corrélation entre l'âge et les intentions de vote. Mais on

sait que le taux de refus des personnes âgées est très élevé. Néanmoins, la comparaison restera correcte si l'on peut admettre que les personnes qui refusent de répondre à l'enquête votent de façon semblable aux autres. Si, en revanche, on a des raisons de penser que ce n'est pas le cas, la comparaison entre des groupes d'âge où les taux de refus sont différents nous induira en erreur. Si dans deux groupes d'âge, par exemple 21-30 ans et 31-40 ans, le taux de refus est le même, la comparaison restera valable, même si le fait de refuser est lié à la position politique.

Si ce qu'on veut faire c'est une estimation des intentions de vote, et même si l'on n'a pas de raison de croire à une corrélation entre refus et position politique, l'échantillon sera biaisé si les taux de refus sont différents aux différents âges. Les personnes âgées, qui votent différemment des plus jeunes, seront sous-représentées, et n'auront donc pas dans l'échantillon le poids qu'elles devraient avoir.

Ces remarques peuvent se systématiser dans le tableau suivant, qui représente une situation où l'on veut à la fois estimer la valeur d'une variable X (dans notre exemple : les intentions de vote), et étudier sa relation avec une autre variable Y (l'âge). Les refus peuvent être ou non liés à X ou à Y.

Taux de refus en corrélation avec X*

Taux de refus en corrélation avec Y**		OUI	NON
	OUI	Comparaison biaisée Estimation biaisée de X et de Y	Comparaison correcte Estimation de X correcte Estimation de Y biaisée
	NON	Comparaison correcte Estimation de Y correcte Estimation de X biaisée	Comparaison correcte Estimation correcte

* Dans notre exemple : les personnes qui refusent de répondre votent différemment (ou non) de celles qui acceptent.

** Dans notre exemple : le taux de refus est lié (ou n'est pas lié) à l'âge.

Cet exemple est lui-même simplifié, car il est fréquent qu'on soit obligé de faire intervenir plusieurs variables explicatives simultanément. Mais le principe du raisonnement est toujours le même, et s'applique à tous les types de biais systématiques.

Il ne faut pas confondre, comme on le fait parfois, les erreurs d'échantillonnage avec les biais systématiques dont il est question ici. L'erreur d'échantillonnage est liée au fait

qu'on ne prend qu'un échantillon et non la population totale. Un écart est donc toujours possible entre les caractéristiques de l'un et de l'autre. Cet écart, on sait en estimer la valeur probable lorsque l'échantillon a été obtenu par une méthode aléatoire. C'est lui qu'on appelle l'erreur d'échantillonnage dont le risque est inévitable. Sa caractéristique est de ne pas être systématique : si on prélève un autre échantillon dans les mêmes conditions, l'erreur pourra être de grandeur différente, de signe contraire. Toutes les techniques d'inférence statistique, qu'il s'agisse d'estimation de grandeurs ou de test d'hypothèses, en tiennent compte. Cette erreur peut être réduite en augmentant la taille de l'échantillon (cf. 2.2.6.).

Les différents biais dont nous avons parlé, au contraire, sont systématiques : ils faussent les résultats toujours dans la même direction. Comme ils ne sont pas aléatoires, les techniques statistiques les ignorent. Il ne faut pas compter sur elles pour les corriger, et on ne peut pas les réduire en augmentant la taille de l'échantillon. Il est donc illusoire d'espérer « améliorer » un échantillon simplement en interrogeant plus de personnes. Si le procédé d'échantillonnage reste le même, les biais systématiques ne changeront pas. La taille d'un échantillon n'est pas une mesure suffisante de sa qualité.

2.2.5. L'échantillonnage pour les enquêtes non statistiques

Lorsqu'on utilise des méthodes non standardisées, entretiens non directifs ou entretiens structurés, il est inutile d'interroger un très grand nombre de sujets. La lourdeur de l'analyse rend difficile l'exploitation systématique d'un nombre important d'entretiens. D'ailleurs l'expérience montre que, pour les thèmes habituellement abordés par ces méthodes, il est rare qu'on voie apparaître des informations nouvelles après le vingtième ou le trentième interview.

En fait, le nombre de sujets nécessaires dépend non seulement de l'hétérogénéité des réactions dans la population face au problème posé, mais aussi, et surtout, de la méthode d'analyse employée, et de l'utilisation qu'on entend faire des résultats. Si l'on attend des entretiens libres principalement un recensement de thèmes, une typologie et des indications sur le vocabulaire utilisé, en vue de la conception d'une enquête systématique par questionnaire, vingt entretiens suffiront en général largement, les suivants ne venant que confirmer ce qu'aura apporté l'analyse des premiers. Mais si on procède à une analyse de contenu plus raffinée, tenant compte par exemple des relations entre les thèmes et de leur succession dans le discours, on pourra à la limite trouver des éléments nouveaux dans chaque interview supplémentaire, quel qu'en soit le nombre. Cependant, dans tous les cas, il arrive un moment où le rendement devient décroissant.

On est facilement tenté à ce moment de faire de l'inférence statistique implicite, constatant par exemple qu'un thème est apparu chez plus de personnes qu'un autre, qu'il apparaît plus souvent dans une catégorie de sujets que dans une autre, que deux thèmes sont souvent associés, etc. Si des hypothèses peuvent être suggérées par ce type d'inférences (tout est bon pour élaborer des hypothèses, dès lors qu'on les reconnaît comme telles et qu'on ne les confond pas avec des conclusions), il faut bien voir qu'il n'y a rien de plus, et qu'on ne se trouve pas dans des conditions où de telles inférences sont, en toute rigueur, possibles. Chacun des sujets interrogés s'étant exprimé à sa ma-

nière, ils ne sont pas suffisamment comparables entre eux pour qu'on admette, sans plus, que les dénombrements ou d'autres mesures qu'on peut faire sur chaque entretien sont comparables. On ne se trouve donc pas dans une situation où une inférence de type statistique est légitime.

C'est pourquoi il est inutile de prévoir un nombre important d'interviews : la qualité de la recherche n'en sera que très faiblement améliorée. Il n'est d'ailleurs pas indispensable d'en fixer le nombre à l'avance (bien qu'on soit souvent, en pratique, amené à le faire pour évaluer le budget de l'étude) ; il serait préférable de prendre connaissance des entretiens à mesure qu'ils sont réalisés, et s'arrêter lorsque leur utilité décroît visiblement. Cette manière de faire présente en outre l'avantage de pouvoir modifier la consigne si on s'aperçoit qu'elle n'est pas bien adaptée aux objectifs poursuivis, ou qu'il serait intéressant d'aborder, en plus, d'autres problèmes.

Dans ces conditions, essayer de constituer un échantillon « représentatif » de la population étudiée n'a guère de sens puisque, de toute façon, on ne fera pas d'inférences globales, et que le nombre de sujets interrogés sera faible. Ce qui est important, c'est de s'assurer de la variété des personnes interrogées, et de vérifier qu'aucune situation importante pour le problème traité n'a été omise lors du choix des sujets. On se rapproche donc plus d'un échantillon par quota, chaque catégorie étant également représentée, que d'un échantillon statistiquement représentatif.

Même lorsqu'on a pris garde à couvrir tout le champ de la population étudiée, un biais important reste possible, dû à la plus ou moins grande aisance verbale des différentes personnes interrogées, aisance évidemment liée, entre autres, à la catégorie sociale. Plus une méthode est non directive, plus elle fait appel aux capacités verbales des sujets. On risque, de ce fait, d'obtenir des interviews beaucoup plus riches, en apparence du moins, de la part de personnes cultivées, ce qui, si on n'y prend garde, peut faire accorder à ce qu'elles disent un poids exagéré [1].

2.2.6. La taille de l'échantillon

C'est une évidence que la qualité, la validité des résultats d'une enquête dépendent de la taille de l'échantillon interrogé. Il est évident que si l'échantillon était égal à la population, l'erreur d'échantillonnage serait nulle [2]. Mais cette référence à la population totale est trompeuse : non seulement la saisir entièrement n'est presque jamais possible, mais le gain de précision par rapport à un échantillon de taille plus raisonnable serait

---

1. Un spécialiste des méthodes non directives, à qui nous signalons le problème, a objecté que les entretiens véritablement non directifs permettent de saisir des mécanismes très profonds, communs à toutes les catégories sociales. Il n'y aurait donc pas de biais : les personnes parlant mieux que les autres exprimeraient plus facilement ce que toutes ressentent. L'objection n'est pas négligeable, si effectivement le problème traité est tel qu'une approche de type psychanalytique est justifiée. Mais il ne faut pas oublier que, dans la plupart des recherches psychosociologiques, et a fortiori sociologiques, ce substrat commun, en supposant qu'il existe, est loin de suffire ; les différences entre groupes sont essentielles, et il faut se garder d'adopter une méthode qui, en prétendant aller au-delà de ces différences, risque plutôt de les camoufler.

2. Ce qui ne veut pas dire qu'une enquête exhaustive serait ipso facto irréprochable. Il reste toutes les erreurs, les biais et les incertitudes liées aux réponses données par les sujets interrogés, dont la validité dépend à la fois de la manière dont les questions sont posées, et de la situation d'enquête, en particulier de la relation qui s'établit entre l'enquêteur et celui qu'il interroge.

minime. Pour s'en rendre compte voyons la formule de l'erreur type sur l'estimation d'une proportion π. Elle est

$$\text{e.t.}(\pi) = \pm \sqrt{\frac{\pi (1 - \pi)}{n}}$$

où π est la proportion réelle,
n la taille de l'échantillon,
e.t. (π) est l'écart type de la distribution d'échantillonnage du paramètre π : autrement dit, on a 95 % de chances pour que la valeur observée sur un échantillon de taille n soit à l'intérieur de l'intervalle :

$$\pm 1{,}96 \text{ e.t. } (\pi)$$

Si la vraie valeur de π est de 0,4, l'intervalle de confiance sera :

	Erreur probable	Intervalle de confiance
pour n = 100	± 0,1	0,3 — 0,5
pour n = 400	± 0,049	0,351 — 0,449
pour n = 1 000	± 0,035	0,365 — 0,435

On voit qu'un échantillon de 1 000 personnes, tirées d'une population qu'on peut considérer comme infinie, permet d'estimer une proportion avec une erreur relativement faible. Ce qui est important surtout, c'est que l'*erreur ne diminue que comme la racine du nombre de sujets* : pour réduire de moitié l'intervalle de confiance il est nécessaire de quadrupler la taille de l'échantillon.

Le nombre de personnes à interroger dépend donc de la précision souhaitée, ce qui n'est pas toujours facile à indiquer. Lorsque les résultats doivent servir à des décisions bien déterminées, il est en principe possible, à partir d'une estimation du coût des différentes erreurs, de calculer la taille optimum de l'échantillon. Mais des situations aussi claires sont rares, et le plus souvent on ne dispose guère de tels guides.

En pratique, d'ailleurs, on ne s'intéresse jamais uniquement à l'estimation d'un paramètre sur l'ensemble de l'échantillon. On prévoit des dépouillements plus ou moins complexes où, par exemple, on cherche à estimer la valeur de ces paramètres sur des sous-échantillons plus ou moins restreints. Il faut donc déterminer la taille de l'échantillon total de façon telle que les estimations effectuées à partir de ces sous-échantillons soient acceptables, ce qui exige qu'on ait déjà une idée, même grossière, de l'importance de ces sous-échantillons.

On dispose encore de moins d'éléments pour déterminer la taille d'un échantillon nécessaire pour mettre en évidence des relations, donc vérifier des hypothèses. Le risque de première espèce étant traditionnellement fixé (à 5 % en général), augmenter le nombre d'observations diminue le risque de seconde espèce, le risque d'accepter à tort l'hypothèse nulle ; autrement dit, lorsqu'on augmente la taille de l'échantillon, on a plus de chances de déceler des relations, même faibles. Comme on ne connaît pas à l'avance l'importance des relations dont on veut vérifier l'existence, prendre un plus grand nombre de sujets revient à augmenter la sensibilité des tests qu'on utilisera.

Mais, ce faisant, on risque de rendre le test sensible à de très petites différences, qu'on peut juger sans intérêt réel.

Faute de critères précis, la meilleure manière de déterminer la taille de l'échantillon consiste à écrire à l'avance les tableaux des résultats, en y introduisant des valeurs vraisemblables, et en voyant si les dépouillements les plus complexes qu'on a l'intention de faire resteront possibles à partir du nombre de sujets prévus. S'ils ne le sont pas parce qu'il y aurait trop peu de sujets dans certaines cases, il faut augmenter le nombre d'observations, ou, si le budget ne le permet pas, bien expliciter les analyses auxquelles on devra renoncer et les limites des analyses possibles.

Dans bien des cas, plutôt que d'augmenter simplement la taille de l'échantillon, ce qui peut amener à accroître le nombre de personnes appartenant à des catégories déjà suffisantes, il y a avantage à constituer un échantillon expérimental, conçu de façon précise en fonction des analyses prévues, ce qui permet d'éviter des échantillons inutilement grands. Mais lorsqu'on change ainsi de stratégie, il faut se demander si tout ce qui avait été prévu reste possible. Par exemple, nous l'avons vu, un échantillon expérimental rend difficile toute inférence directe à la population totale. C'est donc tout le plan de dépouillement qu'il faut revoir.

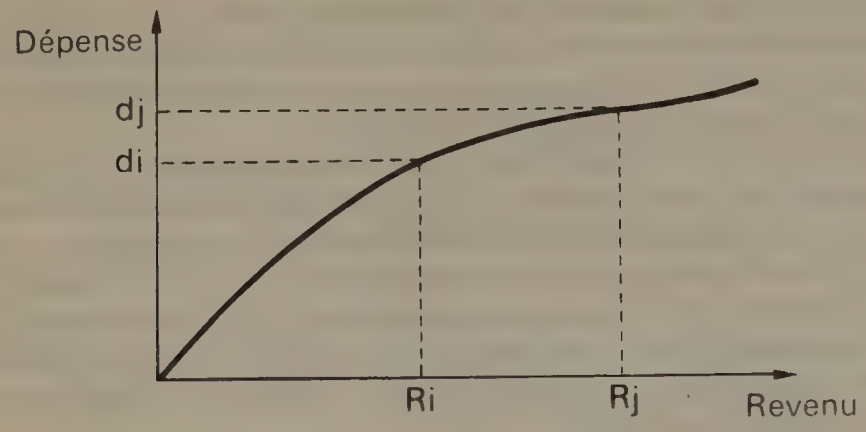
2.2.7. Est-il vraiment nécessaire qu'un échantillon soit représentatif ?

Il est rare qu'on travaille avec un échantillon parfaitement représentatif. D'une part, nous l'avons vu, il y a presque toujours des biais dans la réalisation de l'échantillon, qui n'est pas exactement celui qui avait été prévu. D'autre part, un échantillon représentatif de la population étudiée peut être peu commode, par exemple parce que certains groupes seraient insuffisamment représentés, ou parce que certaines relations seraient difficiles à mettre en évidence. Se poser le problème de la représentativité en soi, et vouloir à tout prix un échantillon parfaitement représentatif, c'est s'imposer une contrainte difficile à satisfaire, et souvent inutile. A la notion globale de représentativité, il faut substituer une notion plus large, celle d'adéquation de l'échantillon aux buts poursuivis, sachant qu'une enquête vise en général plusieurs objectifs (pratiquement, cela signifie qu'il est prévu plusieurs types d'analyse), et que ce n'est pas nécessairement le même échantillon qui, initialement, serait optimum pour chacun. Certains compromis sont alors nécessaires.

Nous avons vu, dans le chapitre consacré aux biais de réalisation de l'échantillon, que le problème se pose différemment selon qu'on veut estimer une grandeur ou vérifier l'existence d'une relation, et selon que le biais supposé est en corrélation ou non avec les variables sur lesquelles porte l'analyse. On peut étendre ces conclusions à tous les écarts à la représentativité, quelle que soit leur origine.

Lorsque l'objectif d'une enquête est d'estimer des grandeurs, la représentativité stricte de l'échantillon est une condition nécessaire de la validité du résultat. Ce point est évident : si, par exemple, nous voulons estimer la moyenne des sommes dépensées par les Français pour un certain poste de leur budget, une proportion trop forte de ménages riches dans l'échantillon entraînera une surestimation de cette moyenne.

En revanche, la contrainte de représentativité est beaucoup moins stricte lorsqu'on cherche à vérifier des hypothèses portant sur des relations. Reprenant l'exemple précédent, supposons qu'on se demande comment varie la dépense en question en fonction du revenu, ce qui se traduit par la courbe de la figure suivante :



La courbe est constituée par les moyennes des distributions de D pour chacune des valeurs de R. La valeur di ne sera pas biaisée du fait qu'il y a, dans l'échantillon, une plus forte proportion de ménages ayant le revenu Ri que dans la population parente.

Les limites de confiance seront plus étroites ; le point sera donc connu avec une précision meilleure, mais on n'aura pas introduit de biais systématique. Tout se passe comme si nous avions constitué indépendamment des échantillons de personnes dans chaque catégorie de revenus.

Le fait qu'il ne soit pas indispensable que la distribution des revenus dans l'échantillon reproduise celle de la population ne doit pas toutefois amener à conclure qu'on peut sans risque interroger n'importe qui, qu'on peut se libérer sans plus de la contrainte de représentativité de l'échantillon. La distribution des revenus dans l'échantillon peut ne pas reproduire celle de la population, mais il reste indispensable que les échantillons de chaque catégorie de revenus soient, eux, représentatifs. Si, par exemple, on constitue l'échantillon sur les lieux où se fait la dépense étudiée, on ne saisira que ceux qui dépensent effectivement pour ce poste, et d'autant plus facilement que leur dépense sera plus fréquente. La dépense moyenne estimée pour chaque catégorie de revenus aura donc toutes les chances d'être surestimée, et comme rien ne nous assure que ce biais sera le même dans toutes les catégories, la relation entre le revenu et la dépense sera faussée.

Il faut également se souvenir qu'une enquête est rarement conçue pour étudier une seule relation. Or, une distribution non représentative des revenus va entraîner le fait que les distributions d'autres variables qui lui sont liées ne seront pas non plus représentatives. on se retrouve devant le même problème : si c'est la distribution dans son ensemble qui est traitée, ce n'est pas grave ; en revanche, si cela entraîne un biais à l'intérieur de chaque catégorie, cela entraînera une déformation de la relation. Il faut réfléchir dans chaque cas particulier aux conséquences possibles de telles distorsions,

en tenant compte de tout ce qu'on peut savoir, ou supposer, sur les relations ou les interactions entre les variables considérées.

Dans certains cas, on peut être amené, pour des raisons de commodité ou d'économie, à se limiter à une sous-population, alors que ce qui est visé c'est la population totale. Par exemple, voulant étudier la population urbaine française, on pourra être contraint de ne faire l'enquête que dans une ou deux agglomérations, au lieu d'en prendre un échantillon. Bien que ce type de généralisation soit très fréquent, on ne dispose guère d'éléments pour juger de sa légitimité, et des bais ainsi introduits. Le fait que, par exemple, la distribution des revenus varie d'une ville à l'autre, n'est pas nécessairement gênant lorsqu'il s'agit de mettre en évidence des relations : le raisonnement que nous avons fait ci-dessus reste valable. En revanche, il faut se demander quelles sont les particularités locales qui risquent de fausser les résultats, même sous la forme de relations. Par exemple, dans les villes proches de Paris on observe que les personnes les plus aisées s'y rendent fréquemment, pour leurs achats ou pour leurs loisirs, ce qui pourra affecter fortement les relations des divers comportements avec le revenu, surtout s'il s'agit d'une enquête sur les lieux de dépenses, et non à domicile. Les conclusions ne seront pas généralisables à des villes plus éloignées de Paris, et de ce fait plus autonomes.

On rencontre des problèmes du même type si on limite l'enquête à une catégorie sociale particulière. Même si celle-ci est définie de façon large de telle sorte qu'on y rencontre une gamme très étendue de revenus, il se peut que la relation observée soit propre à cette catégorie. Par exemple si, comme cela arrive pour un certain nombre de consommations de loisirs, une certaine dépense est plus liée au niveau culturel qu'au revenu lui-même, on risque de ne pas observer de corrélation entre cette dépense et le revenu dans un groupe homogène du point de vue du niveau culturel, alors qu'on l'aurait observée dans une population culturellement plus hétérogène, même avec la même étendue des revenus. Selon ce qu'on vise, on pourra dire que c'est l'absence de corrélation dans un groupe homogène qui est « vraie » : c'est le niveau culturel, et non le revenu, qui est le facteur vraiment *explicatif* des différences de dépenses. Mais on peut aussi estimer que, au niveau *descriptif*, savoir que dans l'ensemble de la population cette dépense varie avec le revenu constitue une information utile.

Cet exemple peut paraître artificiel. Il est en effet peu probable qu'on cherche à généraliser à toute une population à partir d'un groupe homogène du point de vue d'une variable aussi importante que le niveau culturel. Mais les mêmes problèmes peuvent se poser pour des variables beaucoup moins bien définies. Il se peut, par exemple, que certains biais de réalisation de l'échantillon ne se manifestent directement sur aucune des variables étudiées, mais que l'échantillon constitué soit très biaisé selon d'autres variables, non identifiées, qui pourraient jouer le même rôle que le niveau culturel dans notre exemple. Un tel risque, à peu près impossible à évaluer, constitue à lui seul une raison importante de veiller aux procédures de constitution de l'échantillon.

BIBLIOGRAPHIE

En français, le principal ouvrage sur les méthodes de sondage est :

- J. DESABIE, *Théorie et pratique des sondages*, Paris, Dunod, 1966.

On pourra aussi consulter :

- M. REUCHLIN, *Les Méthodes statistiques*, Paris, P.U.F., 1976.

On trouvera aussi de nombreuses indications pratiques dans l'ouvrage de

- MOSER et KALTON, mentionné au chapitre 1.

Chapitre 3

COMMENT INTERROGER ? LES ENTRETIENS

3.1. LES DIFFÉRENTS MODES D'INTERROGATION

A la question « comment interroger ? » répondent des techniques très diverses, qu'il est commode de classer selon la plus ou moins grande liberté laissée à la personne enquêtée de choisir le point de vue selon lequel elle répondra, les thèmes qu'elle abordera et les termes qu'elle utilisera. Pour employer un langage plus précis, les techniques d'interrogation diffèreront selon leur plus ou moins grande *directivité*.

A une extrémité de cette dimension, on trouvera l'*entretien non directif* (appelé aussi *libre* ou en *profondeur*) au cours duquel on propose un thème que le sujet développe à sa guise, les interventions de l'interviewer se bornant à des relances ou des encouragements, sans apporter aucune information ni orientation nouvelle. A l'opposé, le *questionnaire fermé* est constitué par une suite de questions dont l'ordre et la formulation sont fixés à l'avance, et où le sujet qui donne sa réponse doit choisir dans une liste, elle aussi constituée à l'avance, celle qui lui convient le mieux. Entre ces deux extrêmes, de nombreux intermédiaires sont possibles.

Il est d'usage de réserver les termes d'*entretien* ou d'*interview* (qui sont en général synonymes) pour les techniques les moins directives, et d'appeler *questionnaire* les formes d'interrogation où les questions sont formulées à l'avance. Mais, pratiquement, il n'y a pas consensus sur les limites de chacun de ces termes.

Pour définir cette classification, nous avons parlé de la *liberté de la personne interrogée*, et pas, comme on le fait parfois, de celle de l'enquêteur. En effet c'est, à notre sens, une erreur de croire que les méthodes non directives laissent une plus grande latitude à l'enquêteur. Les règles en sont très strictes ; en revanche ces méthodes recourent fortement à son jugement ; les risques d'erreur sont donc plus grands que dans les méthodes plus directives.

Dans la variété des techniques qu'on peut imaginer, nous en préciserons quatre qui jalonnent bien cette dimension de directivité croissante :

1. *L'entretien non directif* : l'enquêteur propose un thème et n'intervient que pour relancer et encourager.

2. *L'entretien semi-directif* : (appelé aussi parfois *clinique* ou *structuré*) : l'enquêteur connaît tous les thèmes sur lesquels il doit obtenir les réactions de l'enquêté, mais l'ordre et la manière dont il les introduira sont laissés à son jugement, la consigne de départ étant seule fixée.

3. *Le questionnaire ouvert* : la formulation et l'ordre des questions sont fixés, mais le sujet peut répondre aussi longuement qu'il le désire, et il peut y être incité par des relances de l'enquêteur.

4. *Le questionnaire fermé* : la formulation des questions, leur ordre et la gamme des réponses possibles sont fixés.

En pratique, il est fréquent de recourir à des méthodes mixtes, comme, par exemple, un entretien non directif immédiatement suivi d'une phase semi-directive, ou un questionnaire alternant questions ouvertes et questions fermées.

L'adaptation aux particularités de l'enquêté, qui est un des avantages des techniques peu directives, peut aussi être, dans une certaine mesure, assurée par le recours à un questionnaire en « arbre » où le choix de certaines questions est déterminé par les réponses du sujet aux questions précédentes, ce qui permet de poser à chacun les questions pertinentes et elles seules.

En résumé, la dimension de directivité croissante, de l'entretien non directif au questionnaire fermé, se traduit par une *programmation* croissante du comportement de l'enquêteur d'abord, et parmi les techniques où celui-ci est déterminé (questionnaires), par une limitation croissante de la gamme de réponses possibles pour le sujet.

Ceci dit, deux caractéristiques restent communes à toutes les formes d'interrogation. La première, qu'on peut considérer comme la définition même de la situation de questionnement, a été formulée dès 1924 par Bingham et Moore : « L'interview est une conversation avec un but. »

La seconde caractéristique commune porte sur la situation d'interrogation. En s'inspirant de Rogers, on peut dire qu'un entretien ou la passation d'un questionnaire est une rencontre interpersonnelle se déroulant dans un cadre et dans une situation sociale donnés, telle que soit impliquée la présence d'un professionnel et d'un sujet naïf.

Malgré ces points communs, il convient de ne pas s'attendre à une quelconque théorie dans ce domaine. La réflexion est située au niveau d'une méthodologie et d'une pratique, même si quelques essais timides de théorisation partielle ont eu lieu.

3.2. CARACTÉRISTIQUES GÉNÉRALES DE L'ENTRETIEN

La définition de l'entretien comme une conversation ayant un but présente l'avantage d'être suffisamment large pour englober une grande variété d'entretiens possibles, mais en contrepartie reste trop vague pour permettre de distinguer les différents types d'entretiens.

La distinction essentielle devant, à notre sens, être faite entre l'entretien à but de diagnostic (qu'il soit par conséquent thérapeutique ou psychiatrique, cette seconde catégorie ne se confondant pas totalement avec la première, ou qu'il s'agisse, par exemple, d'un entretien de sélection) et l'entretien « d'étude » qui vise une généralisation.

Premier type : l'entretien est situé dans le contexte général d'une relation d'aide. L'enquêteur (E. - symbole que nous utiliserons par la suite pour le désigner) étant non pas tellement un professionnel de l'entretien qu'un thérapeute (au sens large du terme) qui utilise l'entretien dans le but de soigner. C'est-à-dire, en fait, dans le but de modifier le comportement actuel de la personne, considéré comme anormal dans le sens primitif du terme, hors norme. Dans ce cas, l'E connaît les méthodes nécessaires pour rendre le comportement normal (adapté !) et utilise l'entretien soit comme un de ces moyens (entretien thérapeutique), soit comme une source d'information (entretien psychiatrique) qui permettra de déterminer certains symptômes et partant de porter un diagnostic. Ce dernier type d'entretien serait faussement assimilé à un entretien d'étude (ou de recherche), dans la mesure où le psychiatre a une grille de lecture, de décodage, du langage du patient, et où, par conséquent, le langage du patient est considéré dans sa globalité comme univoque. A telle déformation langagière correspond telle lecture.

Par ailleurs, il est à noter que la relation praticien-patient est dissymétrique et peut, très généralement, être située dans une relation du type « aide demandée - aide accordée ».

Enfin, cet entretien ne suppose (ou ne devrait supposer) aucun contrôle de la forme, de l'étendue et du contenu de la réponse puisque celle-ci seule, dans toute sa complexité, est garante de la validité du diagnostic porté ou de l'intérêt libératoire de l'acte de parole. Le traitement approprié est celui qui s'applique au « cas » et, à ce titre, ne permet aucune généralisation complète, tout au plus une assimilation à d'autres « cas ».

Second type : l'entretien est demandé par l'E. et présente donc un ensemble de caractéristiques différentes de celles qui précèdent.

On peut, dans un premier temps de l'analyse, les situer dans l'un des quatre niveaux suivants :

— *Contrôle* d'un point donné à des fins de validation partielle de résultats obtenus par ailleurs (c'est le cas d'entretiens post-expérimentaux, afin de vérifier, par exemple, la vraisemblance de la situation expérimentale) ; dans ce cas, l'entretien n'est pas la méthode essentielle.

— *Vérification* d'un domaine de recherche dont on connaît déjà la structure, mais dont, par exemple, on veut voir si certains facteurs ont évolué.

— *Approfondissement* d'un champ dont on connaît les thèmes essentiels mais qu'on n'estime pas suffisamment défriché dans tel ou tel de ses aspects.

— *Exploration* d'un domaine que l'on ne connaît pas.

Les rôles de l'E. et de l'enquêté (e.) sont très différenciés sans que l'on puisse nécessairement parler d'une relation dissymétrique. Dans la plupart des cas l'E. n'a en principe d'autre statut que celui de professionnel de l'enquête (même pas de l'entretien) et ne saurait, par conséquent, instaurer sur ce seul fait une dissymétrie en sa faveur [1]. Par ailleurs, c'est lui qui est demandeur.

Les entretiens d'étude n'ont pas (ou peu, en mettant les choses au pire) de conséquences sur le comportement ultérieur des e., quand bien même l'e. donnerait des renseignements intimes sur ses attitudes, expériences diverses, etc. (l'E. n'a aucun pouvoir sur lui).

L'e. ne se soumet pas à l'entretien pour répondre à un quelconque besoin d'aide thérapeutique, mais bien parce qu'il accepte à un moment donné (pour des raisons diverses et vraisemblablement — sauf cas exceptionnel — peu importantes) de soustraire une partie de son temps à ses activités normales. L'e. n'attend pas, non plus, un quelconque bénéfice personnel de l'entretien auquel il s'est soumis de plus ou moins bonne grâce.

L'E. (ou le chercheur) a des hypothèses (ou devrait en avoir) lorsqu'il fait un entretien, non une sémiologie, et par conséquent ne peut prétendre traiter les résultats à partir d'une « théorie » préexistante (ou pour être plus simple, d'une classification). Le langage, dans ce cas, conserve toute son ambiguïté, la relation sociale toute sa complexité, au niveau des interactions E.-e., et les influences réciproques E. ⇄ e. sont reconnues et analysées. Ce qui n'est apparemment pas le cas dans le cadre psychiatrique. « Ils (les infirmiers) ne considèrent pas celui-ci (le malade) comme un être humain simplement ordinaire - il est, au départ, un « malade » dont il faut mettre en évidence les symptômes... le comportement de l'interné est de toute façon et a priori décrit et interprété en termes de pathologie » (Gentis).

Suivant le but de l'étude, des méthodologies allant d'un contrôle absolu de la forme, de l'étendue et du contenu de la réponse (questions fermées) à une absence de contrôle (théorique, malheureusement !) de ces trois termes (entretien non directif) sont possibles.

Enfin, le traitement des données conduit à des agrégations des réponses fournies aux différents entretiens, à des « procédures de réduction et d'inférence... à des additions pour décrire ou expliquer les faits sociaux ».

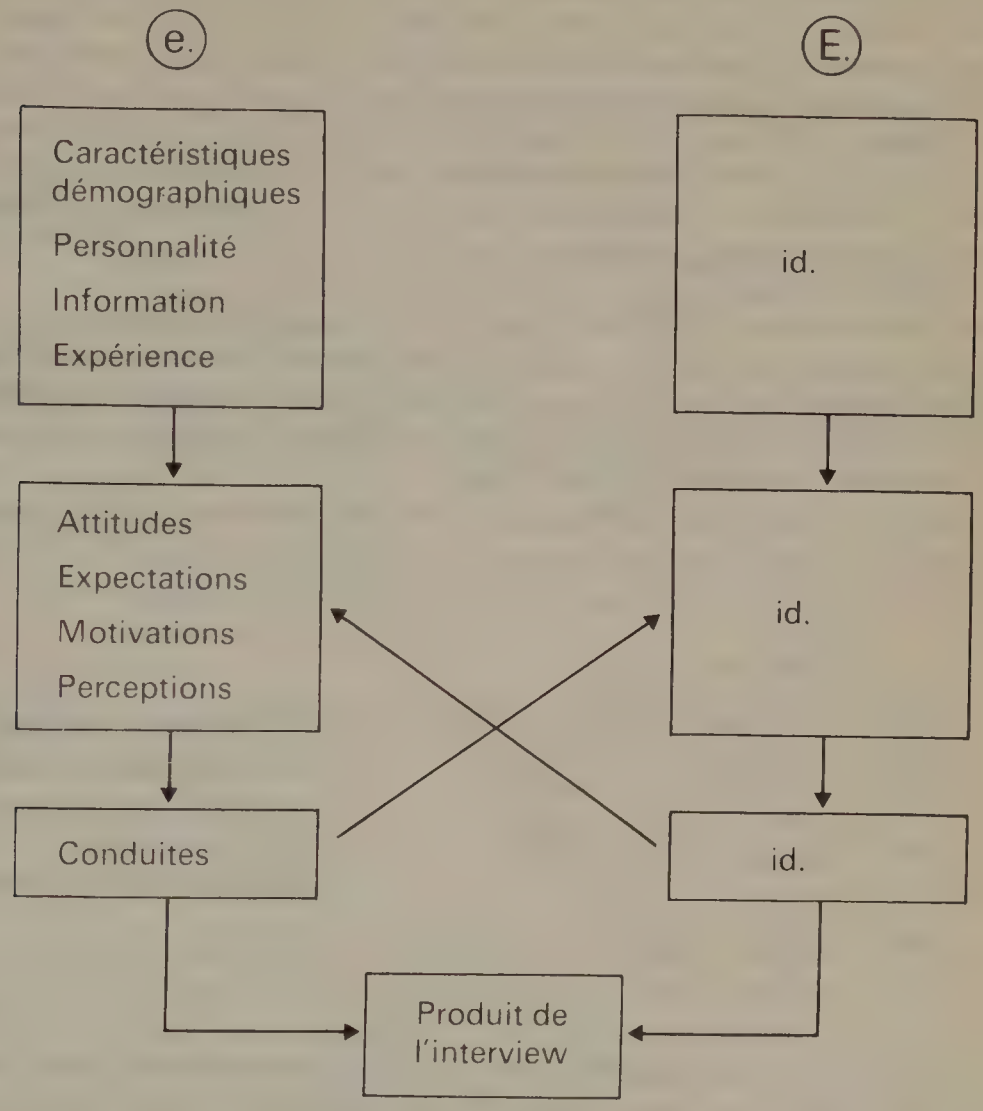
On ne parlera dans cet ouvrage que des entretiens relevant du second cas. Notre compétence et notre croyance dans les autres entretiens étant toutes deux limitées.

---

1. En fait, la situation n'est pas toujours aussi claire et certaines personnes projettent sur l'enquêteur diverses représentations et, par exemple, essaient de transformer la relation d'enquête en une relation d'aide. C'est à l'enquêteur à définir la situation sans ambiguïté, et aussi à résister, pendant la durée de l'enquête, à sa propre tendance à venir en aide.

3.3. LA SITUATION D'ENTRETIEN : CONCEPTS MIS EN JEU

Au terme d'une analyse mettant l'accent sur trois concepts clés : l'accessibilité, les conditions cognitives et la motivation, Kahn et Cannell proposent un modèle motivationnel « qui traite la situation d'interview comme un processus social et l'interview comme un produit social ».



Dans ce modèle, l'accessiblité renvoie au recueil de l'information et à ses obstacles : oubli, répression (au sens analytique), problèmes langagiers, de vocabulaire, etc. Les conditions cognitives renvoient à l'existence de cadres de référence connus. La motivation à répondre est liée à différents facteurs décrits par différents chercheurs : l'autorité légitime, les motivations intrinsèque et instrumentale, l'altruisme, les satisfactions émotionnelles et intellectuelles.

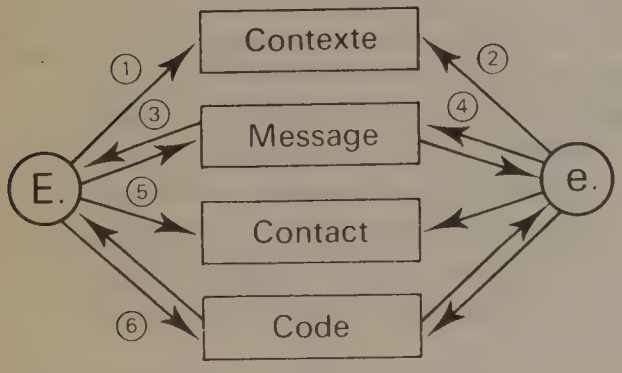
Mais un tel modèle ne nous semble pas rendre compte de façon exhaustive de l'ensemble de la situation d'entretien d'une part — ne serait-ce que parce qu'il oublie très précisément la situation spécifique d'entretien (dans la rue, dans un café, en groupe, en dyade, etc.) — et parce qu'il ne donne pas l'ensemble des interactions possibles, ni ne les indique au sein du rapport E.-e., d'autre part.

Un modèle plus approprié nous semble être celui de la communication adaptée aux processus d'influence. Dire cela suppose un choix théorique : il n'existe pas de non-directivité absolue. En d'autres termes, si le questionnement est lié au contrôle de la forme, de l'étendue, et du contenu de la réponse, et si ce contrôle peut être total (question fermée) ou inexistant (entretien non directif), on prétendra que le continuum n'est pas aussi étendu qu'on veut généralement bien le dire et que la non-directivité est un leurre si on la croit capable d'abolir le contrôle. Tout au plus peut-on dire que l'on est situé dans un schéma E.-e. d'influences réciproques tel que la relation E.-e. suit un processus d'influence supposant l'existence :

— d'un effet à distance,

— d'un véhicule de cet effet, dans le cas précis le langage, qui présente la double caractéristique de signifier et d'être socialisé. Il signifie car il est intentionnel (tout au moins dans le cas précis d'un processus d'influence, située dans un champ sémantique cette notion prendrait une autre valeur), il est socialisé dans la mesure où il est partagé par un ensemble de personnes et leur sert à agir.

Ceci conduit à proposer le modèle suivant, pour rendre compte de la situation d'entretien :



1 et 2 renvoient à la situation,
3 et 4 renvoient aux différentes caractéristiques de l'interviewé et de l'interviewer,
5 et 6 renvoient au langage.

Chacun des termes composant ce modèle sera analysé indépendamment des autres pour des raisons bien compréhensibles de clarté. Cependant, il ne faut pas penser que ceci traduit la réalité d'un entretien où les différentes interactions sont imbriquées et où il apparaît souvent difficile de connaître avec précision les déterminants et leurs poids respectifs. Mais ceci rend compte de l'inadéquation du concept de contrôle et du leurre qu'il constitue en ce qui concerne l'entretien. En effet, dans le cadre d'un entretien, le sujet n'est pas « libre » de dire ce qu'il veut dans la mesure même où il est déterminé par les éléments composant ce qu'il est convenu d'appeler la situation d'entretien.

En fait, si l'on simplifie le schéma en ne retenant comme lieu d'analyse que l'e.,

c'est-à-dire le lieu de production du discours, but de la recherche, on aboutit à proposer cinq schémas rendant compte du modèle précédent et de ses implications :

1 E. — M — e.

Facteurs agissant liés à la situation : récompense, lieu où se déroule l'entretien, temps dont dispose e., institut de recherche effectuant l'enquête, type d'enquête, etc.

2 E. — M — e.

Facteurs agissant liés aux caractéristiques de e. lui-même : mémoire, motivation, etc.

3 E. — M — e.

Facteurs agissant liés aux caractéristiques de e. appartenance à un groupe ethnique déterminé, aspect physique, etc.

4 E. — M — e.

Facteurs agissant liés à E., mais médiatisés par le message : valence du vocabulaire utilisé, style, clarté du message, etc.

5 E. — M — e.

Facteurs agissant liés au message (M), dans la mesure où celui-ci est bien compris, où il évoque quelque chose, où les reformulations sont judicieuses, etc.

Voici maintenant, pour chacun des facteurs, quelques points qui nous paraissent importants.

3.3.1. Facteurs liés à la situation

Lié à la situation on peut noter le *lieu* où se situe l'entretien. Il n'est pas, en effet, indifférent que l'entretien ait lieu dans un endroit calme ou bruyant, dans un bureau ou dans la rue, là où l'e. travaille ou habite. Il n'est pas davantage indifférent que l'E. aille le trouver, ou au contraire que l'e. se rende dans le bureau de l'E. Cet ensemble de lieux peut avoir des conséquences quant au statut et au rôle que l'e. s'attribue à lui-même par rapport au regard que l'E. porte sur lui. Ainsi, l'entretien qui se déroule dans la salle du Conseil de l'O.R.T.F. (par manque de place ailleurs vraisemblablement) avec, en début d'entretien, un porte-mine et autres gadgets offerts à l'interviewé, et en cours d'entretien un rafraîchissement apporté par une « gracieuse » hôtesse, n'aura vraisemblablement pas la même valeur pour l'e. qu'un entretien au thème identique fait dans la rue, au milieu des passants, du bruit des voitures, etc. Partant, il faut essayer d'accorder le but de l'étude et les lieux d'entretien, dans toute la mesure du possible. Ainsi, et à titre

d'exemple uniquement, il ne semble guère judicieux de réaliser des entretiens sur la qualité des émissions de télévision dans un grand magasin, lieu dans lequel les gens se rendent en général dans un but précis, ni en famille car il devient très rapidement difficile de démêler les influences des membres de la famille sur le discours de l'e. Dans ce cas précis, un bureau destiné à ce travail, situé à l'O.R.T.F. si possible (pour des raisons de concordance entre l'objet de l'entretien et son lieu), pendant un temps affecté à cette seule interaction, apparaît préférable.

On peut également penser au temps *réel* dont dispose l'e. et, par conséquent, prendre soin d'annoncer, avant, la durée de l'entretien. Ce problème n'est d'ailleurs guère séparable du précédent. Il est tout à fait certain — et on peut ici se hasarder à généraliser — que la situation optimale est dans bien des cas (au moins pour les entretiens non directifs) celle dans laquelle l'e. a été relativement coupé, pendant la période de l'entretien, des activités qui ont précédé et de celles qui suivront. Ceci signifie, par exemple, que s'il peut être profitable d'interroger une ménagère sur « l'acte d'achat de produits alimentaires » alors qu'elle se trouve dans un supermarché, il est quasiment nécessaire d'avoir dans ce même lieu un coin retiré qui permette de s'extraire des facteurs distrayants (musique, annonces, conversations, etc.) inclus dans l'environnement. De même, s'il peut être intéressant d'interroger un membre du syndicat sur son lieu de travail (pour des raisons d'homogénéité entre le thème de l'entretien et le climat de celui-ci), il est préférable de lui demander de condamner sa porte et son téléphone pendant une durée donnée.

L'organisme qui effectue l'enquête, ainsi que ses buts ou l'image que l'e. en a, ne sont pas étrangers — du moins dans un premier temps — aux réactions de l'interviewé et à ses productions.

Enfin, un dernier facteur que nous voudrions signaler est le fait que l'entretien soit réalisé à travers une relation E.-e., ou au contraire, à travers une situation de groupe de type E.-e. 1, e. 2, e. 3. Les mécanismes mis en jeu dans l'un et l'autre cas diffèrent soit de nature, soit d'intensité, soit de complexité. Il suffit de penser, par exemple, aux processus d'influence, aux problèmes classiques dans les groupes (y compris ceux d'entretien) d'émergence de leader, etc., pour se rendre compte que l'on est confronté à deux situations différentes. Dans ce cas, deux types d'approches sont nécessaires pour les analyses : l'une qui s'intéresse aux productions verbales, l'autre qui analyse la dynamique du groupe, et ce afin de tenter de voir quelles sont les interactions réciproques des deux phénomènes.

D'autres facteurs peuvent intervenir dans la situation, mais ils nous semblent d'un intérêt moindre ou d'une fréquence d'apparition trop peu élevée pour qu'il soit vraiment utile d'en parler (récompenses diverses, désir de parler à quelqu'un, etc.).

3.3.2. Facteurs liés à l'interviewé

On peut classiquement les diviser en facteurs d'ordre culturel, mnémonique, cognitif, motivationnel, conjoncturel, physique (mais on ne parlera pas de ces derniers ici car l'E. est demandeur et c'est dans le sens E. → e. que ces facteurs sont déterminants).

Pour le culturel on pense essentiellement au stock verbal de l'interviewé et à sa possibilité de comprendre certaines questions.

On pourrait, dans cet ordre d'idées, citer d'autres facteurs tels que la réticence, la crainte qu'éprouvent certaines personnes à se soumettre à des interviews, le rejet de toute forme d'interrogation extérieure, etc., mais tous ces facteurs nous semblent en fait être liés à deux paramètres essentiels : le langage et la personne de l'E. que nous examinerons par la suite. S'ils composent des caractéristiques culturelles de telle classe sociale ou de tel type d'ethnie, ils ne paraissent nullement être rédhibitoires pour faire des interviews, et nous semblent devoir être rattachés moins à l'enquêté qu'à sa relation à l'E., relation qu'il faut entendre à la fois au sens non verbal du terme (apparence, comportement, attitudes, etc.) qu'au sens verbal (langage utilisé).

Liés à ce type de facteurs, mais distincts d'eux, on peut noter les *facteurs conjoncturels*. En fait, tout ce qui relève de la pertinence du thème par rapport aux préoccupations de l'interviewé. Il paraît en effet parfaitement inutile d'aller interroger les paysans de la Haute-Loire sur les avantages et inconvénients des tours d'habitation que l'on rencontre dans la région parisienne, sauf si le but de l'étude est précisément de connaître les représentations que se font les paysans de la Haute-Loire de ce mode d'habitation. Mais si le but de l'opération est de savoir comment les gens ressentent un tel environnement urbain, comment ils y vivent, etc., le problème n'est pas pertinent. Exemple grossier dira-t-on et type d'erreur « qui tombe sous le sens ». Il faut croire que le bon sens n'est pas la chose du monde la mieux partagée puisqu'on peut voir fréquemment des études qui n'évitent pas ce genre de piège. Par exemple, demander à des populations rurales qui seront déplacées quel est le type d'environnement urbain qu'elles souhaitent, alors qu'elles n'ent ont pratiquement jamais vu et qu'elles n'y ont jamais vécu.

Une autre classe de facteurs conjoncturels est constituée par l'importance du thème pour le sujet au moment de l'enquête. L'évidence veut que le sujet confronté à un thème peu important pour lui (ce qui est différent du cas précédent dans la mesure où le sujet pouvait, à la limite, être impliqué par le thème mais ne rien pouvoir dire de pertinent sur lui) attachera peu d'importance à son discours. A l'autre extrémité, un sujet fortement impliqué par un thème risque d'« ignorer » les expériences déplaisantes, et par là même de tarir les sources d'information.

Enfin une dernière classe de facteurs conjoncturels concernant l'e. est liée à l'ambiguïté ou à la confusion de son rôle d'interviewé. Dans la mesure où l'on estime que la situation d'interview est une situation sociale, c'est-à-dire une situation impliquant des rôles et des statuts différenciés, il est clair qu'un rôle entraînera des dysfonctionnements au niveau de la relation.

Sans entrer ici dans la (ou les) problématique(s) des théories du rôle, on peut toutefois signaler que le rôle de l'e. est une de ses caractéristiques essentielles, que cette caractéristique lui est attribuée par l'E. au début de la relation E.-e. et qu'elle doit, par conséquent, répondre à divers impératifs :

— définition du statut d'e.,
— définition claire du rôle, non-recouvrement des rôles différents,

— définition claire des attentes de rôles,
— non-incompatibilité des exigences attachées au rôle,
— rôle constant durant le temps de l'entretien.

De la même façon, mais on l'abordera plus loin, l'E. doit définir son propre rôle.

Les facteurs mnémoniques sont dans leur énoncé d'une extrême simplicité et sont liés à la capacité qu'a le sujet de se remémorer un certain nombre d'informations liées au thème. Ils obéissent aux lois classiques concernant la transformation du souvenir : réduction, accentuation, dramatisation, etc. Le problème demeure toutefois — dans la mesure où l'étude fait appel de la part de l'e. à des souvenirs — de pouvoir reconstituer la réalité objective. Ainsi, si dans une usine on enquête auprès de l'ensemble du personnel sur les événements les plus importants qui se sont déroulés au cours des cinq dernières années, sur leurs causes et sur leurs effets, et que l'on utilise la méthode des entretiens non directifs, il est vraisemblable qu'on se trouvera confronté au problème de la transformation des événements dans la mémoire au cours du temps. Dans ce cas, il est quasiment indispensable d'avoir une relation (écrite, par exemple, à travers des chiffres, des comptes rendus du comité d'entreprise et autres) exacte des événements réels. De plus, cela peut apparaître méthodologiquement fructueux (voir comment, en fonction de quelles idéologies, le souvenir a évolué). Par contre, s'il s'agit de connaître l'avis des enquêtés sur le bain moussant, le problème du souvenir est beaûcoup moins crucial.

Les cognitions et affects de l'e., bien que l'on ait habituellement tendance à séparer ces deux types de concepts et les phénomènes auxquels ils renvoient, mettent en évidence, à notre sens, deux types de comportements (ou plus exactement, nous classons sous cette rubrique deux types de comportements) :

— les cognitions qui renvoient au cadre de référence de l'e., à son langage conceptuel, à la connaissance du thème, etc.,

— les affects qu'on peut relier aux perceptions de e. concernant tant la relation présente E.-e., que l'ensemble des relations e.-x (x étant une entité quelconque, individu, chose, événement) réactivées par le thème.

Par ailleurs les comportements issus de ces considérations sont des caractéristiques de e. dans la mesure où ils sont reliés à l'expérience personnelle du sujet.

En ce qui concerne le cadre de référence du sujet, on peut trouver des exemples, du plus simple au plus complexe, pour démontrer son importance (depuis, par exemple, la valence du mot « ordre moral » pour un fasciste et un communiste, en caricaturant du compliment à l'injure, jusqu'à l'analyse faite par Violette Morin des articles de *France-Soir* concernant les délits « du larcin au hold-up » où elle démontre qu'il existe toute une hiérarchie dans l'admiration ou la haine du chroniqueur. Hiérarchie allant d'une certaine admiration pour le vol audacieux commis en plein jour et sans victime, jusqu'à un dégoût cracheur pour les agressions nocturnes de vieilles dames seules. Cadre de référence qui institue une aristocratie et une plèbe du hold-up.).

Le cadre de référence du sujet étant ce qui détermine les connotations des mots uti-

lisés, donc ce qui en partie en définit le sens, il est primordial de le connaître le mieux possible [1].

Pour ce qui est des perceptions du sujet, elles sont également reliables dans la plupart des cas à son cadre de référence. En d'autres termes, l'individu interviewé a forgé (et continue d'ailleurs), au cours de son existence, un certain cadre de référence qui est sa grille de lecture de l'événement. Que celui-ci soit entendu, lu, vu, n'a alors — on en conviendra — qu'une importance secondaire. Ce qui conduit à distinguer la situation et ce qu'il en a perçu. Et, parallèlement, à ne plus séparer ses cognitions et ses affects dans la mesure où tous deux sont des réactions orientées en fonction d'un déterminisme commun.

Il existe, bien entendu, des facteurs cognitifs ne rentrant pas dans ce cadre. Par exemple : e. sait-il quelque chose du thème de l'interview proposé ? Est-il à même de le comprendre ? (encore qu'ici la formule devienne ambiguë). Mais on dira que cela relève le plus souvent de méthodes d'échantillonnage (et des erreurs éventuelles) que de caractéristiques. Le fait que le thème n'évoque rien pour e. peut être interprété comme la caractéristique « il ne sait rien du sujet proposé », mais peut également être perçu comme une caractéristique des chercheurs « ils ne savent pas constituer un échantillon répondant au but de leur étude ». Entre ces deux ignorances, à qui attribuer la caractéristique ?

Enfin, restent les *facteurs motivationnels* qui, par le jeu de la littérature consacrée à ce sujet et des observations des différents praticiens, se sont transformés en quasi-traits de personnalité.

Pour de nombreux auteurs, notamment américains, la motivation de e. pour répondre à l'interview est à rechercher dans le besoin de maintenir à un niveau acceptable par lui « l'estime de soi » (*self esteem*). Par le désir d'être perçu par E. comme une personne ne violant pas les normes sociales et présentant une image flatteuse : celle de la consistance cognitive, du mérite et de la connaissance, sans pour cela qu'un petit brin d'originalité soit mal venu dans cet ensemble de traits. Ainsi Wenar, relate que les mères biaisent l'histoire du développement de leur enfant pour le faire paraître plus précoce ; Lamale conclut, après avoir analysé les comportements des consommateurs relatés par les interviews, que ces comportements sont pertinents pour ce qui concerne l'automobile, les soins médicaux, etc., mais faux pour ce qui est des achats de boissons alcooliques (alors qu'il est notoire que les Américains sont de grands buveurs de whisky [2] !). Tout simplement, ils ne veulent pas le dire par peur de baisser dans leur *self esteem*. Clark et Wallin comparant les interviews des maris et de leurs femmes respectives en ce qui concerne la fréquence des relations sexuelles, trouvent une forte corrélation pour les partenaires exprimant de la satisfaction, mais une faible corrélation lorsque les partenaires sont insatisfaits.

Quoi qu'il en soit, même si le *self esteem* est un facteur important, il existe, dans la

---

1. Au niveau de la formulation des questions ou des thèmes d'un interview directif, le problème n'est plus alors « comprend-t-il ? », mais « que comprend-t-il ? ». Ceci conduit à poser différemment le problème du pré-questionnaire et de sa méthodologie. On y reviendra plus loin.

2. Ceci n'est qu'une note des auteurs du présent ouvrage.

motivation à répondre ou à ne pas répondre, un grand nombre d'autres variables, dont certaines peuvent être considérées comme des caractéristiques des sujets (de par leur fréquence empirique exclusivement et non à cause d'une quelconque théorie de la personnalité). Il existe cependant deux ensembles de variables : l'un lié à la situation (acceptation du thème, compréhension, etc.), l'autre lié à E. (aspect physique, technique d'interview, etc.) qui interviennent dans la motivation de e. sans pour autant constituer des caractéristiques personnelles.

3.3.3. Facteurs liés à l'enquêteur

Les facteurs que nous isolerons ici sont, pour une partie, parallèles à ceux décrits précédemment à propos de e. On peut représenter l'affirmation qui précède de la façon suivante :

• Niveau culturel • Niveau intellectuel • Niveau d'intérêt par rapport au thème etc.	• Langage utilisé • Concepts verbalisés etc.	• Aspect physique • Niveau culturel • Opinions personnelles • Techniques utilisées etc.
↑ Ensemble des facteurs propres à e.	↑ Ensemble des facteurs propres à E. et à e. mais intervenant au niveau de la possibilité de communication à travers le véhicule de celle-ci : le langage	↑ Ensemble des facteurs propres à E.

Un premier type de variables est celui concernant les caractéristiques physiques de E.. Son sexe, son âge, son appartenance apparente à une race, à une classe sociale donnée, etc., ont un rôle certain dans le déroulement de l'entretien.

Ainsi, une enquête auprès de camionneurs menée par des femmes permettait d'obtenir d'excellents interviews (bien que l'on ne prétende pas que ce soit là le seul facteur agissant, ni même le principal) ; il en était de même pour une enquête menée par des jeunes filles auprès de vendeurs de fruits des Halles. Il n'est nullement dans notre propos de signifier que cette procédure est généralisable et qu'il faut envoyer de frêles jeunes filles pour réussir les interviews d'hommes qui font des métiers à valence fortement masculine. Chaque étude est un cas d'espèce et demande à être traitée comme telle.

Concernant l'âge, on pourrait citer la grande difficulté que rencontrent les adultes à interviewer de façon satisfaisante les adolescents. Une étude portant sur la perception que les adolescents avaient d'eux-mêmes et de leurs camarades, menée à l'aide d'entretiens semi-directifs, d'une part par des enquêteurs adultes, d'autre part par des enquê-

teurs pouvant être assimilés à des adolescents (étudiants), laissait apparaître une grande distorsion entre les interviews. Ceux produits par les étudiants étant beaucoup plus riches en informations que les autres, sans que l'on puisse pour autant accuser l'homogénéité des groupes interrogés, puisque l'échantillon était issu d'une population homogène et réparti au hasard entre les enquêteurs.

D'autres exemples pourraient être trouvés dans les entretiens (plus ou moins non directifs) faits entre une assistante sociale (à la fois représentante des adultes, d'une institution, etc. ; là n'est pas notre propos) et un délinquant juvénile.

L'appartenance apparente (à travers des caractéristiques physiques telles que l'habillement, le véhicule, etc.) à une classe sociale peut également être considéré comme un facteur attribuable à E.. Faire des interviews avec des ouvriers, dont le lieu de travail habituel est l'atelier, dans un bureau insonorisé, climatisé, « néonisé », avec des vêtements comportant « nécessairement » une veste et une cravate, revient automatiquement à se faire « étiqueter » comme appartenant à la classe des « cols blancs ». C'est, par conséquent, susciter chez e. des comportements et des productions liés à cette perception.

Il en est de même pour ce qui concerne l'appartenance apparente à un groupe ethnique déterminé. Une étude faite aux Etats-Unis, ayant pour but d'interviewer des vendeurs de biens sur les attitudes envers les juifs, montre qu'il y a des différences significatives dans les entretiens obtenus par des E. au type sémite marqué et les entretiens réalisés par des E. ne présentant aucune caractéristique sémite. Les auteurs, pour obtenir ces résultats, divisent les E. en trois groupes :

1. Ceux qui sont sémites de nom et d'apparence.
2. Ceux qui sont sémites de nom et pas d'apparence.
3. Ceux qui ne sont sémites ni de nom ni d'apparence.

Les trois ensembles d'entretiens correspondant aux trois groupes d'E. montrent des différences importantes, notamment au niveau de l'expression d'attitudes anti-sémites ou de stéréotypes anti-juifs. Différence qui va dans le sens d'une élévation de ce type d'attitude des groupes 1 à 3.

On trouverait également des études concernant les Noirs américains et montrant sans ambiguïté que les différences raciales entre E. et e. (E. noir, e. blanc, ou inversement), conduisent à des résultats différents de ceux que l'on obtient lorsque E. et e. appartiennent à la même race.

Un deuxième type de variables concerne ce que nous avons appelé précédemment le cadre de référence, concept qui peut être considéré comme le produit de l'éducation, des expériences, des convictions morales, religieuses, politiques, etc.

De nombreuses études ont mis en évidence le fait que le cadre de référence de l'E. (et ses corrélats : attitudes et comportements) jouait un rôle dans les résultats de l'interview (on entend par-là dans la production langagière).

Le problème est ancien puisque Rice, en 1929, mit en évidence le fait. Au cours d'une étude sur les causes de la déchéance des individus, menée auprès de personnes transitant dans des hôtels bon marché, Rice trouva deux réponses principales :

— l'une référait la déchéance à un penchant « vers la bouteille »,

— l'autre référait la déchéance aux conditions industrielles et au blâme social ; or, le premier type de réponse correspondait à un enquêteur, le second à un autre enquêteur.

Poursuivant ses investigations Rice, suspicieux, découvrait que le premier E. était un prohibitionniste (acharné) et le second un socialiste (non moins acharné). Ceci met en évidence : à la fois le fait que le cadre de référence (les opinions) de l'E. joue un rôle certain, et le fait que la non-directivité (si elle existe !) ne peut résulter que de la connaissance et de la maîtrise du plus grand nombre possible de facteurs intervenant dans la situation.

Un troisième type de variables concerne la compétence technique de E.. Celle-ci doit, en principe, conduire à surmonter les inconvénients liés à des caractéristiques personnelles difficilement modifiables. Il est bien évident qu'il n'est pas souhaitable d'accumuler des impossibilités telles que la meilleure technique possible ne pourra les résoudre (un Noir des « Black Panthers » interviewé par un Blanc du « Ku Klux Klan » sur le problème racial aux Etats-Unis, par exemple). Mais, sous réserve de prendre garde aux facteurs déjà évoqués, la technique peut permettre de pallier pour une part certaines caractéristiques.

On peut contester le fait de placer la compétence technique parmi les variables attachées à l'enquêteur. En fait, c'est là un parti pris répondant à l'idée exprimée précédemment, et selon laquelle l'entretien implique la présence d'un « professionnel » et d'un sujet. Idée qui, poussée jusqu'à ses extrêmes conséquences, conduit à tenter de couler l'enquêteur dans le rôle défini par l'étude, rôle qui implique des comportements techniquement déterminés pour produire le meilleur résultat possible. Quelle est, dans ce schéma qui risque d'être mécaniste, la part de l'individu ? Nous ne saurions le dire. Notons tout au plus qu'elle existe et qu'elle fait la différence entre deux bons techniciens de l'entretien.

L'ensemble des facteurs liés à la situation, à l'enquêté et à l'enquêteur, a permis de mettre en évidence une notion cruciale : celle de rôle. Il existe en fait deux types de rôles :

— L'ensemble des rôles habituels que joue l'individu et qui sont liés à tout ce qu'on a pu voir (ses croyances, sa race, son milieu socio-économique, etc.).

— Le rôle que l'individu joue lors de la situation d'entretien.

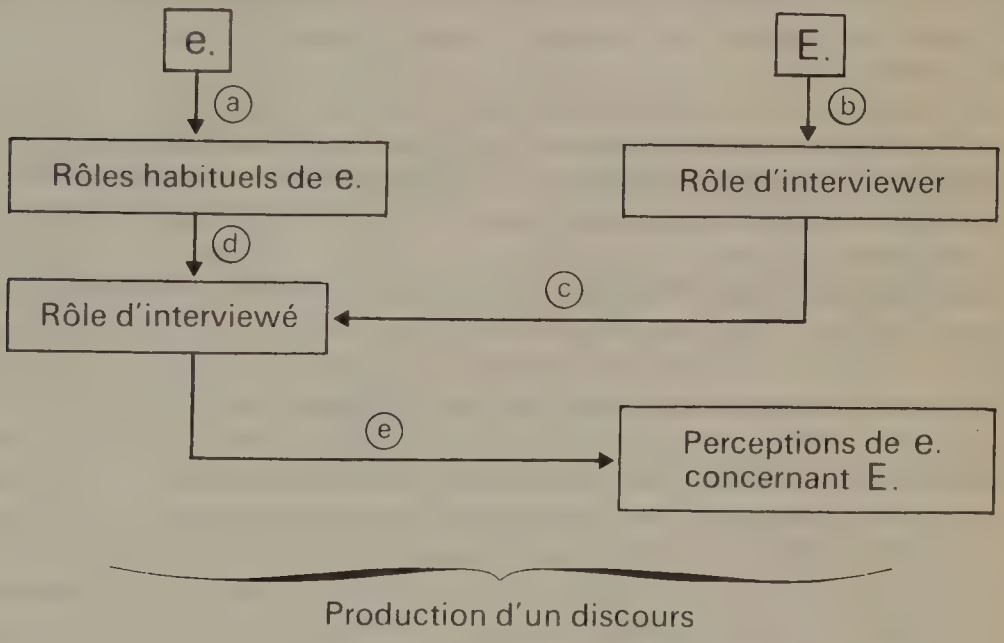
Cette dichotomie conduit à s'interroger sur les relations existant entre ces deux types de rôles, ou plus exactement sur les statuts respectifs de ces rôles.

On se trouve confronté, dans une moindre mesure, aux problèmes liés à l'expérimentation. C'est-à-dire qu'ici, comme là, l'individu est extrait de son « environnement » normal et plongé dans une situation dont les caractéristiques sont l'anormalité, tant au point de vue de la durée de l'interaction que du caractère exceptionnel de la situation. Dans le cas de l'entretien, existe cependant une différence fondamentale avec la situation expérimentale : le stimulus - qui est le thème - conduit l'individu à « se raconter » d'une façon ou d'une autre. Différence qui accentue encore l'ambiguïté d'une situation où la prise de rôle consiste à raconter à un professionnel (dont c'est le seul rôle) qui écoute, les expériences, affects, attitudes, etc., liés aux rôles habituels.

En d'autres termes, on serait assez tenté de dire que l'on a d'un côté un ensemble de rôles que l'individu interrogé ne peut quitter comme on se défait d'un vêtement et que, de l'autre côté, on lui impose un nouveau rôle qui s'ajoute aux autres mais ne les élimine pas. De là l'irruption, dans le rôle actuel, des rôles habituels et les éventuels conflits de rôles (se manifestant sous la forme de refus divers, par exemple refus de se souvenir, refus de parler au-delà d'un certain seuil, etc.) qui en résultent.

Par ailleurs, l'E. a un rôle et un statut qui lui est attaché. Mais là encore il y a interférence entre le rôle joué par l'E., son statut d'E., et les perceptions de l'e. Tout se passe en quelque sorte comme si E. ne présentait pas une image univoque et que la plurivocité des indices soit interprétée par e. pour situer E. hors de ses rôles et statuts actuels. Le conflit de rôles se situe ici davantage au niveau du rôle présenté (par E.) et des rôles induits (par e.) qu'au niveau de E. seul. Même si l'on supposait que E. se confonde complètement avec son rôle d'enquêteur (c'est-à-dire une personne à l'écoute d'une autre, utilisant certaines techniques pour faciliter l'expression de cette autre, mais n'imposant en aucune façon ses propres opinions), on ne pourrait pas (ou mal) maîtriser le phénomène des perceptions de e., qui biaisent nécessairement la relation.

D'une façon générale, la relation E.-e. apparaît être une relation de rôles, telle que :



3.3.4. Facteurs liés au langage

On entend par langage, dans le cas précis, le message allant de E. → e. en ce qu'il correspond à une production linguistique. A ce titre, les facteurs qui lui sont liés sont de deux ordres :

a. Ceux qui constituent des indices permettant à e. de situer E..

b. Ceux qui relèvent du message lui-même, c'est-à-dire de la possibilité de le comprendre, d'y répondre, etc.

Il est clair toutefois que cette distinction n'est pas aussi tranchée que ce qu'il pourrait apparaître à la lecture de ce qui précède et que, par exemple, un message compréhensible constitue également un indice que e. peut interpréter pour situer socialement E., et réciproquement le style utilisé par E., qui est un indice (ou peut l'être) de son statut socio-économique, peut également faciliter ou non la compréhension.

On voudrait cependant insister sur cette différence, à notre sens fondamentale. Le message sert à situer E. d'une part, mais est en tant que tel véhicule de signification et produit à ce titre un effet chez le récepteur (e.).

Dans la première classe de facteurs (a), on trouve des variables telles que le style utilisé, la clarté du message, les mots employés, etc. Dans la deuxième classe de facteurs (b), on trouvera la possibilité de comprendre le message, c'est-à-dire en fait la possibilité de le décoder donc de faire référence à un code commun. Ceci renvoie aussi bien au sens habituel du terme lui-même (celui qu'on trouve dans les dictionnaires), qu'à l'univers des connotations que e. attribue aux termes employés.

Ces deux classes de facteurs apparaissent par conséquent comme le produit des variables précédentes, telles que le cadre de référence, les préoccupations de E. et de e., de leurs niveaux culturels, mais ces facteurs ont aussi leur signification intrinsèque.

Le langage doit donc :

— Être accessible (à e. essentiellement dans le cas d'un entretien non directif, et à E. et e. dans le cas d'entretiens semi-directifs ou directifs), c'est-à-dire qu'il est inutile de demander à une personne peu scolarisée ce qu'elle pense de « l'introduction du moi transcendantal dans la littérature paupérisante de la fin du XIX[e] siècle » (aux personnes scolarisées non plus !).

— Permettre une réponse. En effet, l'e. peut comprendre la question, c'est-à-dire le langage utilisé, mais celle-ci renvoie à une organisation conceptuelle inhabituelle, par exemple « votre femme est-elle introvertie ou extravertie ? » ; chaque mot peut avoir une signification précise pour e., mais celui-ci n'a jamais pensé à sa femme en ces termes.

— Motiver l'e. à répondre, par exemple en incluant dans le thème la notion de validité de toutes les réponses possibles, ou de tous les discours possibles.

— Être conforme aux attentes de rôle de e. concernant E., c'est-à-dire le plus proche possible de son propre univers langagier, partant de sa perception et de sa verbalisation des événements, et en définitive de son cadre de référence. D'où quelques problèmes lorsque l'enquêteur bien intentionné veut faire comprendre à l'ouvrier qu'il attend de lui une production sur ses conditions d'existence en atelier et lui demande « quel est votre vécu au niveau de la chaîne ? ».

Chaque mot est compris bien souvent, la réponse est possible, la motivation à répondre peut être forte, mais l'ouvrier interrogé, n'attendant pas que E. se comporte comme un intellectuel ouvriériste, le renverra poliment mais fermement à ses chères études, si éloignées d'une chaîne que la relation de rôle (le jeu de rôle) est impossible, car invraisemblable.

Cette dernière notion de vraisemblance est tout à fait essentielle pour que la relation

E.-e. ait quelque chance de fonctionner de façon satisfaisante. Et la vraisemblance passe obligatoirement (et sous réserve d'incompatibilités non rédhibitoires par ailleurs) par l'utilisation d'un langage adéquat.

La situation d'entretien nous apparaît par conséquent assimilable à un jeu de rôle (au sens expérimental et non thérapeutique) : c'est-à-dire à la création d'une situation artificielle dans laquelle sont distribués des rôles inhabituels. Le problème étant non pas le caractère artificiel de la situation, concept non pertinent pour notre objet, mais la vraisemblance de la situation ; celle-ci repose sur :

— l'adéquation de l'E. à son rôle d'E., c'est-à-dire non pas sa neutralité, mais sa capacité à approcher le cadre de référence de l'e. et à s'y conformer dans ses attitudes, comportements, et surtout dans son langage.

— la clarté de la définition du rôle de e., et sa motivation à le jouer (sa capacité étant liée à la compétence de E. pour une grande part).

— la possibilité de maîtriser les conflits de rôles nés de la concomitance de rôles habituels et du rôle actuel d'une part (ceci est essentiellement vrai pour e.), ou de la non-conformité entre les conduites en rôle théoriques de E. et les perceptions que e. a de E. (comportement, attitudes, etc.).

3.4. LES DIFFÉRENTES SORTES D'ENTRETIENS

Parmi les objectifs des entretiens libres, il est utile de bien distinguer deux catégories :

— Obtenir des informations d'ordre *cognitif* : comment le sujet organise-t-il le champ qui lui est proposé ? Quelles sont les limites qu'il donne à ce champ ? Quelles relations y perçoit-il ? Quels concepts et quel langage utilise-t-il ? Plus généralement, quelles sont les représentations de la situation et les normes en fonction desquelles il agit ? etc.

— Obtenir des informations d'ordre *affectif* : comment le sujet ressent-il le domaine en question ? Quelles sont ses attitudes ? Quel est son « vécu » à ce propos ?

En revanche, il nous semble parfaitement illusoire de prétendre y voir, comme le font certains de façon plus ou moins publicitaire, une possibilité d'exploitation de l'inconscient. Qu'on puisse, dans certains cas, tenter quelques interprétations de type psychanalytique, ou simplement symbolique, à partir du discours ainsi obtenu, c'est possible, mais il faut bien admettre que ce qu'on obtient est, en général, des plus « superficiel ». Certes, et c'est même un des intérêts majeurs de la méthode, on tire souvent d'un entretien non directif autre chose qu'une reformulation du discours manifeste du sujet (même cela, d'ailleurs, n'est pas à dédaigner), et on est donc amené à tenir un nouveau discours que le sujet lui-même ne reconnaîtrait peut-être pas nécessairement pour sien. Toutefois, ce qu'on saisit ainsi reste en général très loin du refoulé que recherche le psychanalyste ; on atteint plutôt ce que nous proposons d'appeler un *inconscient cognitif* que le sujet n'a pas pu expliciter lui-même, faute de concepts et de théorie adéquats, faute aussi parfois d'y avoir suffisamment réfléchi. Il arrive d'ailleurs fréquemment, au cours d'un entretien non directif bien mené, que le sujet découvre lui-

même l'organisation de son champ, et certains de ses aspects, mais la facilité même avec laquelle il en prend conscience montre qu'il n'avait pas été vraiment refoulé ; plus simplement, il n'avait jamais encore eu l'occasion, ni pris le temps, d'y réfléchir dans des conditions favorables [1].

Ces deux aspects, cognitif et affectif, de l'entretien libre, peuvent être rattachés à deux des principales sources historiques de la méthode, les travaux de Piaget et ceux de Rogers.

Piaget, bien qu'il ne soit presque jamais cité à ce propos, semble avoir été le premier à introduire systématiquement dans une expérimentation rigoureuse, sous le nom de *méthode clinique* (devenue plus tard, plus justement, la *méthode critique*) une attitude de l'expérimentateur qui ne soit pas entièrement standardisée, mais qui cherche à s'adapter à chaque sujet, adoptant, dans toute la mesure du possible, ses concepts et sa représentation de la situation, donc une attitude radicalement opposée à la standardisation des tests ou des procédures expérimentales. La méthode consiste à proposer une tâche bien définie mais suffisamment ouverte, pour que des réponses très éloignées de la pensée du chercheur adulte puissent être données et développées ; dans son exploration ultérieure l'expérimentateur intervient en posant de nouveaux problèmes, toujours choisis en fonction des réponses précédentes, ou en discutant des réponses données, en montrant les contradictions, non pour faire comprendre, mais pour explorer la structure cognitive sous-jacente, dont on s'efforcera toujours de ne pas sortir, évitant pour cela tout apport de raisonnements nouveaux, si ce n'est très consciemment, à titre de réactif.

La méthode clinique, apparue donc en psychologie de l'enfant au début des années vingt, s'est répandue quelques années plus tard chez les sociologues. Elton Mayo raconte comment, se trouvant devant des comportements, adultes cette fois, qu'il ne comprenait pas, il eut l'idée de transposer la méthode de Piaget, dont il venait de prendre connaissance. Les entretiens effectués dans le cadre de la célèbre expérience de la Général Electric, et qui ont été à l'origine du mouvement des « relations humaines » dans l'industrie, ont été conçus et menés selon cette méthode.

L'attitude « critique » de l'expérimentateur piagétien, qui intervient en faisant approfondir une réponse, qui pose de nouveaux problèmes, et qui ne craint pas de mettre le sujet devant ses contradictions est, par ces aspects, évidemment très éloigné du non-directivisme de Rogers. C'est que les problèmes, et les arrière-plans théoriques, à l'origine des deux méthodes, sont profondément différents. Le non-directivisme rogérien est, méthodologiquement, sinon théoriquement, la systématisation de l'attitude d'écoute et de neutralité bienveillante de la psychanalyse freudienne. Comme cette dernière, initialement, elle a été développée dans une intention thérapeutique, et non de recherche. Le premier objectif du psychothérapeute rogérien n'est pas de connaître mais,

---

1. Le fait que nous refusions de voir, dans un entretien non directif, une psychanalyse en miniature, n'implique pas que certains concepts d'origine psychanalytique ne puissent pas nous être utiles. En particulier, dans une situation aussi peu structurée, et où l'on peut être amené à aborder des thèmes affectivement importants, reconnaître l'existence de résistances, de rationalisations, par exemple, est essentiel. De même, analyser le « contre-transfert », à la fois de l'interviewer et de celui qui procède à l'analyse de contenu, peut être utile et mettre en garde contre une pseudo-objectivité. Quant à l'interaction elle-même entre l'interviewer et l'interviewé, elle a fait l'objet d'un certain nombre de recherches, dans des cadres théoriques divers. Nous y reviendrons.

par son attitude, de faire en sorte que son patient se rende véritablement responsable des conclusions auxquelles il aboutit, et des décisions qu'il peut prendre. On trouve aussi chez Rogers l'idée d'une spontanéité vraie qui ne peut s'exprimer réellement qu'en supprimant les cadres cognitifs et moraux extérieurs à l'individu.

3.4.1. Différents types d'entretiens

On distingue classiquement trois types d'entretiens :

— Non directifs (ou libres).

— Semi-directifs.

— Directifs ou standardisés (ce dernier type d'entretien est fort proche du questionnaire à questions ouvertes).

Chacun de ces types d'entretien présente des caractéristiques propres, qui interviennent moins au niveau des techniques (au moins en ce qui concerne les entretiens non directifs et semi-directifs) qu'au niveau des thèmes et de leur agencement.

Il n'est pas utile de développer longuement ce point, on dira simplement que :

— Dans le cas d'un entretien non directif, l'E. se contente de poser le thème de l'entretien dont les caractéristiques essentielles sont d'être large et ambigu. La notion d'ambiguïté est fondamentale dans ce type d'entretien puisque c'est elle qui permet à e. de développer sa propre pensée à propos d'un thème très général n'incluant aucun cadre de référence particulier. Et surtout pas celui du chercheur ou de l'enquêteur.

La notion d'ambiguïté doit donc être ici comprise comme la présence d'un thème qui introduit la discussion, mais permet au sujet de l'interpréter à partir de son propre cadre de référence.

— Pour ce qui concerne l'entretien semi-directif, il existe un schéma d'entretien (grille de thèmes, par exemple). Mais l'ordre dans lequel les thèmes peuvent être abordés est libre, si e. n'aborde pas *spontanément* un ou plusieurs des thèmes du schéma, E. doit lui proposer le thème. Dans le cadre de chaque thème les méthodes se rapprochent de celles de l'entretien libre.

L'ambiguïté (définie comme précédemment) est moindre dans la mesure où le schéma d'entretien structure, qu'on le veuille ou non, le sujet, et par conséquent lui impose un cadre de référence. Toutefois, chacun des thèmes du schéma conserve une relative ambiguïté. Ce qui veut dire que les grandes catégories du cadre de référence global étant posées, il demeure une non-définition des cadres de référence au niveau de chaque point (ou catégorie). Ce qui est, par conséquent, défini c'est le champ à travers ses catégories, mais les catégories structurantes demeurent relativement ambiguës.

— Pour l'entretien directif ou standardisé, fort proche d'un questionnaire dans lequel ne figureraient que des questions ouvertes, il n'existe pratiquement plus aucune ambiguïté. L'ensemble du cadre de référence est défini (structure du champ et catégories structurantes) et l'e. doit se situer par rapport à ce cadre, entrer en lui pour pouvoir répondre de façon correcte.

Dans ce cas, les méthodes utilisées par E. sont tout à fait comparables à celles utilisées par un enquêteur classique posant des questions, dans le cadre d'un questionnaire.

Ainsi, ces trois types d'entretiens renvoient dans leur démarche à *l'ambiguïté définie comme l'absence de cadre de référence imposé.*

3.4.2. Quel entretien pour quel problème ?

On distinguera quatre types d'utilisations de l'interview d'étude :

— contrôle,
— vérification,
— approfondissement,
— exploration.

Comme on pourra le remarquer, chacune de ces utilisations présente *un degré de liberté plus ou moins grand* par rapport à un savoir antérieur. Degré de liberté nul lorsqu'il s'agit de contrôle, puisque dans ce cas le savoir antérieur existe, qu'il a été obtenu en général avec d'autres instruments qu'avec l'entretien (on pense à l'exemple cité de l'expérimentation et du contrôle de sa vraisemblance), et que l'entretien n'est dans ce cas qu'une façon de s'assurer que les méthodes précédemment utilisées ne sont pas biaisées d'une façon ou d'une autre.

En d'autres termes, on passe d'un cadre de référence parfaitement constitué à un cadre de référence que l'on voudrait connaître. Ce parallélisme n'est pas de pure forme, mais constitue véritablement le fil directeur de la pensée des hommes du « concret », ceux qui « font » du « marketing » par exemple, ou de façon plus générale ceux qui ont à prendre des décisions à partir d'études. Et, chose étonnante, il constitue également le fil directeur (non présenté de cette façon, il est vrai) de nombreux textes s'intéressant à la question.

Pour illustrer le premier type de préoccupation on remontera à des conférences faites par le directeur du marketing d'une entreprise de « conseils » importante (par la taille en tout cas, 1 500 ingénieurs) à de faméliques étudiants avides de savoir. Il était dit au cours de cette conférence dont le titre était « Entretiens libres dans les études d'opinions et d'attitudes » (que l'on excuse l'analyse de contenu sauvage qui va suivre) :

— « L'entretien est fait en vue de l'action »,
— « un client pose un problème, ce qui entraîne des contraintes de temps et d'argent » (n'est-ce pas là un préambule éminemment concret ?),
— « l'entretien libre a de l'intérêt :
— pour chercher les termes dans lesquels se pose le problème au public,
— construire des modèles explicatifs des comportements individuels,
— dégager des variables qui rendent compte des variations des comportements et attitudes de la population étudiée, »
— « lorsqu'il s'agit de comprendre, le questionnaire n'est pas satisfaisant, car la question oriente la réponse, la question n'est pas toujours comprise, l'interviewé ne peut répondre à la question car il ne s'analyse pas, car sa culture personnelle ne lui permet pas d'exprimer certains concepts, parce qu'il ne veut pas dire pourquoi il fait certaines choses, etc. »

N'est-ce pas là une façon, différente sans doute, mais claire de dire que l'entretien libre s'applique essentiellement à une analyse exploratoire ?

Les préoccupations des chercheurs sont identiques, on se contentera de citer deux d'entre eux (Cannell et Kahn, 1967) : « La question ouverte est plus appropriée quand l'objectif de la recherche n'est pas seulement de découvrir des attitudes ou attributs du répondant, mais aussi d'apprendre quelque chose... sur la structure de base sur laquelle il a formé son opinion, le cadre de référence à l'intérieur duquel il répond à la question. »

Ici encore, et bien qu'il s'agisse de question ouverte, on peut penser que, pour les auteurs, les recherches exploratoires correspondent à un certain degré de liberté au niveau du stimulus inducteur de la réponse.

Cet ensemble de considérations nous conduit à proposer le tableau [1] suivant, dans lequel les cases marquées nous semblent constituer l'adéquation la meilleure entre un type de recherche et une méthode d'entretien.

ENTRETIEN / RECHERCHE	NON DIRECTIF	SEMI-DIRECTIF	DIRECTIF
Contrôle			●
Vérification		●	●
Approfondissement	●	●	
Exploration	●		

Une telle problématique de l'utilisation des méthodes d'entretien peut se justifier par l'existence préalable ou non d'un savoir antérieur, mais elle suppose également que soient définis avec précision les objectifs des différents types d'entretiens et les raisons du choix d'un type d'entretien donné par rapport à une situation donnée [2].

1. Il est évident que les indications portées dans le tableau ne sauraient concerner tous les cas de figures possibles, mais seulement le plus grand nombre possible. Il n'est, en effet, pas impossible que l'entretien non directif soit approprié - dans une circonstance particulière - à une recherche de type contrôle.

2. Il faut encore distinguer un autre objectif possible qui est le recueil d'informations. L'utilisation par les ethnologues d'« informateurs » est classique et pose donc les problèmes propres aux contacts entre cultures. Mais, même dans sa propre société, le sociologue peut avoir, lui aussi, besoin d'être informé. Qu'il s'agisse, par exemple, de porter un diagnostic sur une organisation ou d'analyser une décision pour connaître les mécanismes du pouvoir, il faut être informé de « ce qui se passe », ou de « ce qui s'est passé ». Cet objectif est ambigu : on sait bien que les récits qu'il obtient sont marqués par la place des interviewés dans le système étudié, donc qu'il ne peut pas accepter sans critique comme des informations « objectives ». Mais ces points de vue font aussi partie de la réalité que l'on cherche à connaître, d'où la difficulté de ce type d'entretien, surtout lorsqu'on interviewe une personne ès qualité (le directeur d'une entreprise par exemple). Son point de vue, avec ce qu'il peut comporter de parti pris et de déformations, nous intéresse comme tel, mais il est aussi parfois le seul à pouvoir fournir certaines informations « objectives » dont le chercheur a besoin. Il faut alors recourir aux méthodes de critique des témoignages, que les historiens connaissent d'ailleurs mieux que les sociologues, en particulier à la confrontation avec d'autres sources, lorsqu'elles existent, des documents en particulier.

L'entretien libre. Il invite le sujet à répondre de façon exhaustive, dans ses propres termes et avec son propre cadre de référence, à une question générale (le thème) caractérisée par son ambiguïté.

A titre d'illustration, on reprendra la conférence précédemment citée. De cette dernière il ressort que :

— Les problèmes que rencontre l'entretien libre sont ceux du langage, puis ceux des contextes sociologique, économique, psychologique, culturel, dans lesquels se pose le problème soulevé par le thème.

— Le but poursuivi par l'entretien libre est d'essayer de comprendre le contexte, les besoins, motivations, exigences, les solutions apportées par e., c'est-à-dire l'univers perçu par le sujet comme présentant des solutions au problème, le niveau auquel cet univers est perçu, ce qu'il contient, etc., la création et la transformation des attitudes liées au problème, les processus de décision que celui-ci entraîne.

Généralisant le cadre tracé ci-dessus (qui est vraisemblablement le cadre de référence du conférencier, homme de marketing, rappelons-le, c'est-à-dire soumis aux impératifs de la rentabilité immédiate), on peut référer le choix de l'entretien non directif à plusieurs raisons :

— Il est approprié pour rechercher la structure du cadre de référence d'un sujet, l'intensité de ses affects concernant un problème.

— Il n'est pas sensible à des différences de langage entre chercheur (et interviewer, en principe) et interviewé, et peut donc être utilisé auprès de n'importe quelle population.

— Il permet de mener une étude sans savoir au préalable quel est le niveau d'information des interviewés sur le problème, et même s'il est souhaitable de l'utiliser lorsqu'on pense que le problème est situé hors du champ d'expérience d'une partie de la population interrogée (et que l'on souhaite malgré tout interroger pour des raisons précises).

— Si la structure cognitive du sujet par rapport aux problèmes n'existe pas, soit de façon précise, soit de façon proche du cadre de référence qu'impliquerait nécessairement un questionnaire, il est préférable de pratiquer l'entretien libre. En effet, dans ce cas la réponse à des questions pourrait être largement biaisée.

— Il peut, le cas échéant, et pour des raisons issues plus ou moins directement de ce qui précède, motiver l'e. à répondre. Par exemple, s'il manque d'information sur le sujet, des questions précises risquent de l'indisposer car elles sont révélatrices de son ignorance, s'il n'a pas structuré son champ de connaissance par rapport au problème posé et peut être gêné pour répondre, donc peu motivé, etc.

Finalement le choix de l'entretien libre comme méthode est lié de façon directe (que ce soit chez le « marketing man » ou dans ce qui précède) à la connaissance antérieure que l'on a de la situation que l'on veut analyser. Situation comprenant au moins deux variables : le problème et la population à laquelle on le pose. Cette constatation renvoie bien aux études de type exploratoire, ou *éventuellement* de type approfondissement dans la mesure où on estime incomplète son information sur un sujet et qu'on recherche des informations sur tel ou tel point particulier de ce domaine (et non sur un ensem-

ble de points, ce qui nous renverrait à l'entretien semi-directif), en le traitant comme une exploration limitée à un aspect du problème.

L'entretien semi-directif se définit par rapport aux remarques faites à propos de l'entretien libre. C'est-à-dire qu'il intervient à mi-chemin entre une connaissance complète et antérieure de la situation par le chercheur, ce qui renvoie à l'entretien directif ou au questionnaire (on peut se demander d'ailleurs s'il existe vraiment une limite entre ces deux méthodes) et une absence de connaissance qui renvoie à l'entretien non directif. Dans ce cas, le sujet est invité à répondre de façon exhaustive, dans ses propres termes et avec son propre cadre de référence, à une question générale (le thème) caractérisée par son ambiguïté. Mais s'il n'aborde pas spontanément un des sous-thèmes que l'enquêteur connaît, celui-ci pose une nouvelle question (le sous-thème) dont la caractéristique n'est plus l'ambiguïté, afin que le sujet puisse produire un discours sur cette partie du cadre de référence du chercheur.

On peut expliciter la différence existant entre l'entretien libre et l'entretien semi-directif en disant que, dans le premier cas, le chercheur n'a aucun cadre de référence antérieur, ou l'oublie volontairement ; dans le second cas, le chercheur a un cadre de référence antérieur, mais ne l'utilise que si le sujet oublie une partie de celui-ci. L'entretien semi-directif est donc approprié pour approfondir un domaine donné, ou vérifier l'évolution d'un domaine déjà connu.

L'entretien directif est un ensemble de questions ouvertes, standardisées et posées dans un ordre immuable à l'ensemble des enquêtés. Ces questions supposent de leur part des réponses relativement courtes et précises, à des questions non ambiguës.

Ainsi, dans ce type d'entretien sont définis non seulement le cadre de référence et le langage le plus approprié à l'analyse du problème, mais également l'ordre de progression le meilleur pour étudier le sujet. Cette dernière assertion suppose que l'on connaisse à la fois les éléments composant le cadre de référence afférant au sujet étudié, mais également sa structure. Ces entretiens supposent enfin que l'on connaisse d'une certaine façon (ne serait-ce qu'au niveau de ses productions langagières ou de sa structure conceptuelle) la population interrogée [1].

L'entretien directif suppose donc que :

— On limite l'objectif de l'étude soit à vérifier un problème dans son ensemble ou en partie, en ayant recueilli par ailleurs et à l'aide de méthodes diverses des informations sur le problème, soit à contrôler un point ou l'autre d'un problème traité par ailleurs (plan expérimental dans lequel sont comprises une situation expérimentale donnée et un questionnaire post-expérimental pour vérifier la vraisemblance de la situation expérimentale auprès des sujets, par exemple).

— Les dimensions que l'on veut étudier (que ce soit des attitudes, des perceptions, un état civil, ou autre) soient connues et nettement précisées.

— L'on connaisse le niveau d'information que la personne interrogée a sur le sujet.

---

1. Il convient cependant de remarquer que c'est dans le cadre d'un questionnaire constitué de questions fermées que le contrôle de la forme, de l'étendue et du contenu de la réponse possible est maximal, et qu'il existe, par conséquent, dans le cas de l'entretien directif une caractéristique importante : celle de pouvoir nuancer sa réponse.

— L'on sache si la structure cognitive des enquêtés est constituée sur ce problème (en supposant que l'information soit suffisante) en termes suffisamment proches de ceux que l'on trouve au niveau de la question.

— L'on connaisse suffisamment le langage des enquêtés pour que les questions posées ne présentent pas d'obstacles au niveau du sens, c'est-à-dire pour que celui-ci soit le plus univoque possible.

3.5. LES TECHNIQUES UTILISÉES LORS DES ENTRETIENS

Les techniques d'entretiens doivent être comprises comme l'ensemble des moyens nécessaires mis en œuvre de façon systématique par E. pour mener à bien un entretien. C'est-à-dire pour faire en sorte que l'enquêté se soit exprimé le plus complètement possible sur le thème qui lui était proposé [1].

Le but général poursuivi par l'enquêteur s'articule autour des points que nous avons précédemment évoqués (cf. supra : facteurs liés à l'enquêté), c'est-à-dire que :

— Le langage utilisé doit être accessible par e., constituer un support ayant un sens pour lui.

— Le thème doit constituer un stimulus pour e., être évocateur de quelque chose et appeler une réponse.

— Les rôles E. et e. doivent être nettement précisés par E.

— e. doit être motivé pour répondre.

— L'information recueillie doit être la plus large possible.

De cet ensemble d'exigences vont découler les différentes règles techniques nécessaires pour mener à bien un entretien. On les exposera dans l'ordre chronologique correspondant au déroulement d'un entretien.

3.5.1. Etablissement de la relation E.-e.

L'ensemble des remarques préliminaires faites par E. auprès de e. a pour but de mettre e. en situation de jeu de rôle en situant les rôles respectifs de E. et e.. E. commence par s'identifier en fournissant son nom, le but de la recherche, l'institut, cabinet, etc., qui la fait. Il explique comment e. a été choisi, c'est-à-dire, en fait, les modalités d'échantillonnage utilisées, il indique quelles seront les modalités de recueil des données (magnétophone, notes écrites, etc.), la durée de l'entretien. Il insiste également sur les règles déontologiques de la profession et notamment sur la notion d'anonymat de e., ou plus précisément sur l'anonymat de son discours.

Par ailleurs, il définit ses attentes vis-à-vis de l'enquêté. Par exemple : « Vous pouvez me dire tout ce qui vous passe par la tête, il n'y a pas de bonne réponse, ce que nous désirons connaître c'est votre avis sur le problème qui est posé, etc. ».

---

1. On parlera essentiellement ici des entretiens non directifs et semi-directifs, les techniques de l'entretien directif sont plus proches de celles du questionnaire que de celle de l'entretien et seront traitées à ce titre dans la partie « questionnaire » de l'ouvrage.

Il s'agit essentiellement, lorsqu'on définit le rôle de e., de bien lui faire comprendre et admettre que l'on ne posera pas de questions. L'enquêté a souvent une attitude qui le conduit à demander que lui soient posées des questions, ce qui se traduit au début d'un entretien par des phrases du type : « Que voulez-vous savoir ?... Posez-moi des questions et je vous répondrai ».

Il est donc indispensable de situer le rôle de e. par opposition à cette attitude habituelle. Habituelle, en effet, car dans la vie courante les relations du type « questions-réponses » sont plus fréquentes que les relations du type « écoutant-parleur ». Au cours de l'entretien le même problème peut se poser et appellera la même attitude de la part de E. [1].

Enfin, l'attitude de E. au début et en cours d'entretien doit être d'écoute attentive, de tentative constante de compréhension de ce qui est dit par l'enquêté, de *non-critique* et de *non-évaluation*. Il y a toujours quelque ridicule à vouloir donner des conseils de posture et d'attitude ou autres fadaises de même acabit aux enquêteurs : « Durant un entretien tenez-vous comme ceci, regardez le sujet dans le blanc des yeux, hochez consciencieusement la tête à intervalles réguliers », tout au plus peut-on conseiller l'attention réelle au discours de l'autre (et non les signes extérieurs de cette attention) et surtout — point crucial — l'absence de toute attitude normative qui se traduirait nécessairement par une évaluation (positive ou négative peu importe) de ce qui est dit.

Ces dernières remarques seront déterminantes dans le choix des « techniques » classiquement utilisées pour mener à bien un entretien non directif ou semi-directif.

3.5.2. La formulation de la consigne

Le choix de la consigne à proposer au sujet au début d'un entretien non directif soulève des problèmes délicats. Paradoxalement on peut dire que le fait qu'elle soit « correctement comprise » n'est pas essentiel, contrairement à ce qui est le cas pour les questionnaires. En effet, ce que l'on cherche dans un entretien non directif, c'est l'ensemble de ce qui est évoqué par un certain terme ou une certaine expression. Il n'y a donc pas de vraie « réponse » ou de « bonne compréhension ». Le fait que le terme utilisé soit compris par certains sujets dans un sens différent de celui auquel s'attendait le chercheur, peut constituer, dans la plupart des cas, une information importante dont il serait dangereux de se priver. Tout au plus, faut-il prendre soin d'éviter certains malentendus grossiers, fondés sur une homonymie par exemple, et qui n'ont évidemment aucun intérêt. En revanche, il faut se poser des problèmes de *pertinence* de la consigne par rapport au problème qu'on veut explorer.

Poser ainsi le problème et admettre que ce qu'on recherche dans un entretien non directif, ce sont des réactions à des *termes*, à des *mots*, c'est se placer d'emblée et explicitement sur le terrain *linguistique*, en excluant tout autre point de vue. Cette attitude du chercheur a le mérite d'être claire et rigoureuse, mais peut aussi se révéler inutilement

1. Pour ce qui est de l'entretien semi-directif, les techniques sont les mêmes ici et dans ce qui suivra. On fera donc l'économie d'une dichotomie inutile dans ce cas précis.

limitative, dans la mesure où le champ qu'on cherche à explorer peut ne pas se résumer en totalité par un petit nombre de termes.

Supposons que l'on veuille étudier la vie de quartier. Une consigne comme « parlez-moi de votre quartier » ou « parlez-moi de la vie de votre quartier » pourra nous donner ce que représente, pour chaque sujet, les mots « mon quartier ». Toutefois il se peut que dans ses représentations de l'espace urbain, ou dans l'organisation spatiale de ses activités, on observe certaines caractéristiques que le sujet lui-même ne rattache pas au mot « quartier », mais qui sont suffisamment liées au voisinage du domicile considéré comme un espace particulier, pour entrer dans le champ qui intéresse le chercheur. La démarche de celui-ci pourra alors être l'opposé de la démarche « linguistique » : il ne s'intéressera pas aux significations du mot « quartier » pour les sujets, mais il définira lui-même, a priori, un certain champ et tentera de l'explorer sans être lié à l'un ou l'autre des termes que les sujets peuvent utiliser. La meilleure consigne sera alors celle qui aura le plus de chances de susciter des réponses se situant à l'intérieur du champ ainsi défini.

Le choix entre une telle consigne « réaliste » et une consigne « linguistique » ou « nominaliste » dépend de la capacité du chercheur à définir a priori son domaine. Dans le cas du « quartier » ce terme recouvre un champ dont il n'est pas trop difficile de cerner les limites. A priori ce qui nous intéresse alors c'est le contenu du champ. Dans d'autres cas, ce sont les limites elles-mêmes du domaine qu'on veut connaître ; on devra alors avoir recours à une consigne linguistique, en étant conscient de ces difficultés.

Par exemple, si l'on veut étudier la représentation du psychologue dans différentes catégories de la population, il n'est probablement pas possible de définir au départ les limites du concept : ce serait préjuger de ce qui constitue précisément l'objet de la recherche. On s'en tiendra donc à la consigne « qu'est-ce qu'un psychologue ? », quitte, s'il le faut, à utiliser en complément d'entretiens non directifs d'autres techniques permettant d'explorer systématiquement certains aspects qu'on n'associe pas à ce terme, mais à propos desquels les réactions des populations considérées nous sont nécessaires.

L'exemple du quartier nous permet aussi de mettre en évidence une autre des difficultés dans la formulation de la consigne, le risque d'imposer involontairement au sujet un découpage de la réalité, un cadre de référence, qui ne sont pas les siens, et qu'on ne retrouvera, au terme de l'analyse, que parce qu'on les aura introduits dans la manière de présenter le thème au sujet.

Rien ne nous assure au départ que la structure de l'espace urbain autour du domicile se fera bien conformément à ce qui est suggéré par le mot quartier ou des expressions jugées équivalentes. En d'autres termes, nous ne pouvons pas être certains qu'au départ tous les sujets distinguent dans la ville une zone proche de leur domicile et à l'égard de laquelle ils auraient des attitudes et des comportements spécifiques. En fait, diverses études sur ce thème ont montré que des différences importantes existent sur ce point, liées en particulier à la catégorie sociale.

Le danger qu'on court avec ce type de consigne c'est qu'elle sera en général suffisamment acceptable pour que très peu de sujets la refusent ou même la critiquent. Le mot « quartier » est compris par tous, même si ce n'est pas de la même façon, et le domaine

en question est suffisamment peu précis pour que la plupart des sujets acceptent, momentanément et pour la durée de l'entretien, d'organiser leur discours en fonction de ce concept, même s'ils ne l'auraient pas fait spontanément. Sans le vouloir, par une telle consigne, on aura imposé un point de vue et malgré tous les efforts de non-directivité qu'on pourra ensuite déployer au cours de l'entretien, le poids de cette induction risque de se faire sentir jusqu'à la fin.

Dans le cas particulier, pour les deux raisons que nous avons mentionnées, une consigne du genre « parlez-moi de la vie ici » laissant le sujet libre de choisir ce qu'il mettra dans « ici » nous semble préférable à toute autre qui mentionnerait le mot « quartier ».

Il y a parfois intérêt à ne pas proposer directement dans la consigne le thème qui fait l'objet de la recherche, mais un thème plus large, ce qui présente plusieurs avantages. Premièrement, cela permet de voir comment le thème apparaît, s'il est mentionné immédiatement ou si, dans le champ large en question, il occupe une position périphérique (il faut toutefois prendre garde à ne pas identifier trop sommairement l'ordre dans lequel les différents thèmes apparaissent et leur importance). On pourra ainsi étudier comment le thème apparaît, quels sont les autres thèmes qui y amènent. Dans certains cas, même son absence dans le discours spontané pourra être significative. Deuxièmement, cette manière de faire permet de saisir des attitudes et des représentations plus générales, le discours du sujet n'étant pas sur ce point influencé par une représentation trop étroite qu'il pourrait se faire des attentes de l'enquêteur.

Dans ces cas-là il pourra être également intéressant de commencer par une consigne très large et une technique tout à fait non directive, puis, lorsqu'on estime qu'on a obtenu tout ce qu'il était possible d'obtenir de cette manière, de proposer, s'il n'est pas apparu spontanément un nouveau thème, plus précis, en terminant éventuellement par des questions proprement dites.

L'ordre inverse, par contre, doit être évité. Après des questions précises, pour lesquelles l'enquêteur se sera contenté de réponses brèves, le sujet conservera probablement la même attitude, s'attendant très vite à de nouvelles questions. D'autre part, cette première phase directive risque d'orienter tous les discours ultérieurs, faisant ainsi perdre un des principaux apports de la non-directivité.

Lorsqu'on réfléchit à la formulation d'une consigne, on est souvent amené à se demander s'il y a avantage à la personnaliser par des expressions telles que « pour vous » ou « selon vous », ou par l'emploi de pronoms et d'adjectifs à la deuxième personne, (vous, votre, etc.), ou s'il ne vaut pas mieux l'éviter et poser la question en des termes plus généraux et impersonnels.

Bien que le problème soit important, il n'est pas possible de proposer de règle générale à ce propos. L'une et l'autre attitudes ont leurs avantages et leurs inconvénients. Une consigne trop impersonnelle risque d'être comprise comme demandant une réponse générale et « objective », et suscitera facilement des stéréotypes et un discours lui-même impersonnel. Mais elle pourra aussi, lorsque le thème touche de trop près, permettre au sujet de s'exprimer plus librement, sans être exagérément bloqué par la conscience qu'il aurait de parler de lui-même. Parlant en termes généraux, il pourra peut-être aussi plus facilement parler de son entourage et de la façon dont lui-même se situe

par rapport à celui-ci de façon générale, on peut dire qu'une consigne « personnalisée » indique au sujet qu'on attend de lui qu'il s'implique dans son discours, qu'il parle de lui-même, avec tous les risques de résistance et de fuite que peut provoquer cette situation.

Pour reprendre un des exemples précédents, celui des entretiens sur la vie de quartier, on pourrait hésiter entre « parlez-moi de *la* vie ici » ou « parlez-moi de *votre* vie ici ». La première consigne, vraisemblablement, incitera à parler de ce qui se trouve dans les environs, de ce qui s'y passe. Cela pourra vite tourner à une description, voire à une énumération (il y a l'épicier, et puis il y a la pharmacie, et puis...) d'un intérêt souvent limité. Les aspects les plus personnels risquent d'être omis, le sujet les éliminant comme non pertinents, surtout si, à tort ou à raison, il se considère comme atypique. En revanche, il contrôlera peut-être un peu moins son discours que s'il parlait explicitement et consciemment de lui-même.

La consigne « personnalisée », elle, risque peut-être d'effrayer certains sujets qui pourront la juger indiscrète. Le souci d'apparaître « dans la norme », dont nous verrons qu'elle est une des principales causes de biais dans les enquêtes, pourra les paralyser ou orienter exagérément leur discours. On risque ainsi d'obtenir aussi des récits d'événements ou d'aspects de la vie du sujet qui ne se trouvent liés à l'environnement que de façon très contingente, au détriment de ce qui nous intéresse de façon plus directe. Tout compte fait, pour un thème comme celui-ci où, pour la majorité des sujets, les résistances risquent de ne pas être trop fortes, une consigne personnalisée sera probablement préférable. Mais on voit que le choix n'est pas évident.

Toutes ces précautions dans la formulation de la consigne peuvent paraître exagérées, tatillonnes et formalistes. Celui qui n'a qu'un début d'expérience d'enquêteur ou d'analyste pourra être à tel point frappé par la diversité des discours obtenus à partir d'une même consigne, que ces précautions et ces raffinements lui sembleront byzantins. Effectivement, lorsqu'on se trouve devant une seule interview, il n'est souvent plus possible d'y retrouver les traces de la consigne. Ce que dit le sujet semble s'articuler de façon autonome, la consigne n'ayant apparemment joué qu'un rôle de signal de départ. Mais ce n'est en général qu'une illusion ; lorsqu'on analyse en détail non pas un seul, mais un nombre suffisant d'interviews réalisées à partir de deux consignes différentes, la comparaison fait apparaître des différences non négligeables.

Enfin, il est à noter que, concernant le thème, plusieurs possibilités peuvent être envisagées quant au support utilisé pour le présenter.

Ce support peut être *verbal*, et c'est ce qui est le plus fréquent. Dans ce cas, le thème se concrétise par une phrase relativement courte et ambiguë. D'ailleurs, ainsi qu'on l'a déjà souligné, cette notion d'ambiguïté est fondamentale (surtout dans l'entretien non directif et dans une moindre mesure dans l'entretien semi-directif) car c'est elle qui conditionne la possibilité pour e. de parler à propos de son propre cadre de référence et non d'entrer dans celui de E. (ou du responsable de l'étude). Il peut également être *écrit*, ou enregistré, soit parce qu'on tient à ce qu'il n'y ait aucun biais d'enquêteur à ce niveau, soit parce que l'énoncé du thème est trop long pour que sa compréhension et sa mémorisation soient aisées pour e., à la simple écoute. Il peut enfin être *graphique* (photo, dessin, etc.). On peut, en effet, partir d'une image ambiguë pour mener l'entre-

tien (non directif ou semi-directif). A propos de cette dernière technique, moins utilisée, on citera deux exemples. Dans une étude concernant les dentistes, après une série d'entretiens non directifs et leur analyse thématique, cinq thèmes principaux avaient été dégagés. Afin de vérifier leur réalité et d'approfondir leur signification, cinq dessins ambigus, reprenant chacun des cinq thèmes, furent réalisés. La seconde série d'entretiens, semi-directifs, fut menée à partir de chacune des planches présentées comme stimulus ; les résultats du discours furent analysés à l'aide d'une méthode dérivée de celle du TAT. Dans une étude se rapportant aux transports aériens étaient insérées, au milieu d'un questionnaire, deux images de type projectif devant donner lieu à un discours de la part des questionnés. Ces discours avaient pour but de préciser un certain nombre de points contenus dans le questionnaire.

3.5.3. La relation E.-e. en cours d'entretien

Dans ce paragraphe on approfondira les techniques utilisées en cours d'entretien. Celles-ci sont de deux sortes, les techniques dites « classiques » que l'on peut définir par leur taux élevé d'utilisation, et des techniques particulières applicables en fonction d'un problème spécifique.

Les techniques « classiques ».

Elles mettent en œuvre différents procédés à partir d'un thème verbal ou non (image, écrit, enregistrement) :

— Utilisation d'expressions brèves qui marquent l'intérêt pour ce qui est dit, la compréhension et le désir de voir le discours se poursuivre. Ces expressions sont du type : je vois, oui, je comprends, j'entends bien, mm-mm, etc. A propos de la dernière locution « mm-mm » on doit souligner son emploi délicat, elle peut très rapidement prendre un caractère artificiel, si elle est utilisée inconsidérément ou à mauvais escient.

— Utilisation du procédé dit du « miroir » ou de l'« écho ». E. répète le mot ou le groupe de mots (mais, en aucun cas, une longue phrase, ou de façon plus générale un fragment de discours relativement important) qui viennent d'être prononcés par e.. Cette technique est généralement considérée comme un encouragement à continuer, à poursuivre et à approfondir l'idée exprimée. Cela peut aussi et accessoirement vouloir exprimer l'attention, la « sympathie » (que l'on définira ici comme une écoute compréhensive, et non pas comme le désir de se rendre sympathique. Cette dernière attitude, sur laquelle nous reviendrons, aboutissant souvent à des réactions négatives de e.).

— Réalisation de synthèses partielles ou « reformulation » d'une partie de discours formant un tout. En général, la reformulatiòn est destinée à permettre un approfondissement des idées exprimées par e.. Si l'on considère que la plupart des interviews nondirectives constituent pour e. un « discours-recherche », et lui permettent de structurer ses cognitions — ou de les créer — à propos d'un thème donné, au fur et à mesure que l'entretien se déroule, on admettra l'importance de la « reformulation ». Celle-ci permet en effet à e. d'aller plus loin dans sa propre analyse du problème. Mais l'intérêt de

cette technique est corrélé avec son danger car au cours de la reformulation, E. peut aisément faire passer (sans nécessairement s'en rendre compte) ses propres conceptions, et ceci d'autant plus aisément qu'il a déjà des hypothèses implicites sur le champ d'étude.

— Formulation de demandes « neutres » (ou le plus neutres possible) d'information additionnelle à propos de discours, demandes qui peuvent se traduire par des formules telles que « que voulez-vous dire ; qu'avez-vous en tête en ce moment ; à quoi pensez-vous... ? » ou encore « pouvez-vous me dire davantage sur cela ? ».

— Formulation de demandes particulières qui peuvent prendre la forme d'une question : « pourquoi pensez-vous cela ; quand cela se passait-il ; c'est-à-dire plus précisément ? ». Mais ce procédé doit nécessairement avoir, indépendamment de son caractère de demande, un caractère de relance d'un discours plus ou moins bloqué. En effet, si ces deux caractères n'apparaissent pas conjointement et que seul l'aspect questionnement existe, on risque fort de réinstaller e. dans une attitude antérieure d'attente de questions, et par là même de bloquer l'entretien.

— Répétition du thème le cas échéant, mais en prenant garde que cela n'acquière pas une connotation négative pour e., essentiellement « vous vous écartez du sujet ».

— Utilisation des silences à des fins de réflexion pour e.. Dans ce domaine il convient de distinguer les silences courts — n'excédant pas 5 secondes — et les silences longs — plus de 10 secondes. Si les premiers ont dans la plupart des cas un effet positif, les seconds ont généralement un effet négatif. Malheureusement les silences longs peuvent s'installer ; dans ce cas il ne reste plus comme solution à E. qu'une capacité à supporter le silence supérieure à celle de e.

Ces différentes techniques n'ont pas une place précise à un moment privilégié de l'entretien. Tout au plus peut-on donner quelques indications très globales (et très sujettes à modification suivant les cas d'espèce) et très fragmentaires.

L'interview peut être décomposée — au niveau du discours — en trois parties :

— Une partie préalable dont la durée est de l'ordre de 10 à 15 minutes, pendant laquelle le discours est plus ou moins un ensemble de stéréotypes concernant le thème, partie dans laquelle on a toutes les chances de retrouver les clichés circulant dans une culture et/ou une sous-culture (celle de la population interrogée) à un moment donné. Au cours de ce premier stade les interventions de l'enquêteur sont, en général, limitées et ressortent de l'utilisation d'expressions brèves.

— A la fin de cette partie, e. a bien souvent l'impression d'avoir « tout dit » et d'une certaine manière il a effectivement tout dit. Plus précisément, il a donné sa connaissance immédiate et stéréotypée du thème. Dès lors commence son « discours-recherche » au cours duquel il faut l'aider à clarifier ses propres idées, à aller plus loin, plus profond, etc. Cette deuxième partie du discours suppose souvent, pour être amorcée, l'usage de la reformulation. Par la suite le « miroir », les « expressions brèves », les « silences » sont utilisés pour que le discours se développe.

— Enfin, une troisième partie est celle où e. ne peut rien dire de plus du sujet et où cela se traduit par une redondance du discours présent avec le discours de la deuxième partie.

Dans cette phase l'utilisation des demandes « neutres », « particulières » ou de la

« reformulation du thème », peuvent permettre d'apporter soit des précisions intéressantes, soit la confirmation de l'épuisement du sujet par e..

Ce découpage du discours et cette correspondance entre les parties du discours ainsi isolées et les techniques à utiliser ne doit pas faire illusion et donner l'impression qu'un entretien non directif ou semi-directif est un mécanisme dont la précision n'a rien à envier au meilleur des plans d'expérience. En fait, ce qui précède ne constitue qu'un ensemble d'indications qui perdraient toute valeur à être utilisée avec rigidité, leur valeur est indicative sans plus. Il appartient à l'interviewer, en fonction de la situation particulière dans laquelle il se trouve, de juger ce qu'il est le plus opportun d'utiliser à tel ou tel moment d'un entretien.

Les techniques « particulières ».

Un certain nombre de techniques liées soit à un choix délibéré du chercheur qui pense qu'une nouvelle façon de faire est intéressante à étudier, soit à une opportunité née de la situation d'étude dans laquelle on se trouve peuvent être dénombrées.

Piaget avait utilisé au cours d'entretiens cliniques avec des enfants, la méthode dite des « *contre-exemples* ». Cette méthode consiste pour l'essentiel à proposer en face des allégations de e. un certain nombre de contre-exemples. Ainsi si e. dit — à propos d'un thème qui serait « qu'est-ce qu'un psychologue ? » — « le psychologue c'est un monsieur qui fait passer des tests pour voir si on est intelligent ou pas », E. lui rétorquera « mais alors la personne qui étudie les relations homme-machine et qu'on appelle psychologue ? ».

La méthode du contre-exemple généralisée à l'ensemble de l'entretien a fait ses preuves avec Piaget, elle demeure cependant délicate à manier car elle risque de bloquer e. en le plaçant dans une situation qui accuse son ignorance. De plus, elle n'est — à notre sens — utilisable que :

— dans un champ dont on connaît déjà un certain nombre de composants objectifs ; pour reprendre l'exemple du psychologue, il est nécessaire de connaître les différentes professions auxquelles conduit la psychologie ;

— dans une population dont on connaît déjà un certain nombre de caractéristiques et notamment son degré d'information à propos du problème. Ainsi si l'on décide de demander « qu'est-ce qu'un psychologue ? » à une population de médecins et à une population rurale, peut-être n'est-il pas souhaitable d'utiliser la même méthode.

En effet, on peut supposer a priori (et notamment du fait que les études médicales comprennent un enseignement de psychologie) que les médecins connaissent l'existence de psychologues et une partie de leur activité, l'utilisation de la méthode du contre-exemple risque dès lors d'être fructueuse, car elle permet d'introduire dans les cognitions déjà structurées de nouvelles informations qui conduiront l'interviewé à restructurer le champ de ses cognitions et à produire par conséquent, du moins peut-on l'espérer, un discours plus riche.

Une modalité parallèle à la précédente est dérivée des théorisations concernant l'influence sociale et consiste à fournir les allégations (réelles ou supposées) d'une source prestigieuse ou dévalorisée. L'utilisation de ce type de technique peut être faite dans un but précis : degré de certitude que e. accorde aux arguments avancés, résistance des

arguments de e. à des sources autres, etc. Ce type de technique convient mieux à une étude de type approfondissement que de type exploration.

Une troisième technique peut être dénommée « *d'incompréhension volontaire* ». Elle demande à E. d'avoir une attitude de non-connaissance du champ à l'aide de phrases telles que « je ne vois pas très bien ce que vous voulez dire, parce que je ne connais pas... vous pouvez m'expliquer davantage parce que je ne connais pas cette profession... ce produit, etc., cette expression signifie quoi ?... ». Utilisée lors d'une enquête effectuée auprès de camionneurs par des enquêtrices, elle a donné d'excellents résultats et des entretiens riches. Devant l'incompétence supposée des enquêtrices (liée au fait vraisemblablement que la relation E.-e. avait la double caractéristique d'être une relation enquêteur-enquêté et homme-femme), légitime en ce qui concerne un métier « d'homme », les enquêtés fournissaient un grand nombre d'informations.

Une dernière technique - au moins en ce qui nous concerne - consiste à utiliser tout ou partie d'une interview déjà réalisée sur le sujet d'étude. Les modalités d'utilisation sont diverses, mais peuvent être classées en trois catégories :

— Réécoute partielle par e. de ce qu'il a dit au cours de l'entretien. Dans ce cas, l'enquêteur centre l'attention sur une partie du discours et sollicite de la part de l'enquêté un approfondissement de celle-ci.

— Réécoute globale par e. de son entretien afin de tenter ensuite d'aller plus loin.

— Réécoute croisée d'interviews. Par exemple, dans le cas de deux populations « parlant » sur le même sujet, on peut vouloir tenter de supprimer les biais liés aux possibilités langagières de chacune des populations, en supposant que l'une, de par son niveau culturel, s'exprime facilement, et l'autre non. Dans cette hypothèse, on fera écouter aux interviewés de la première population un entretien réalisé dans la seconde et, réciproquement, aux interviewés de la seconde population un entretien réalisé dans la première. Ensuite, et ensuite seulement, on effectuera l'entretien avec e.

Enfin, il faut noter, en ce qui concerne les techniques « particulières » qu'elles sont à appliquer avec discernement, c'est-à-dire en fonction d'hypothèses précises quant à ce qu'on recherche, et au meilleur moyen de l'obtenir. Les appliquer aveuglément à toute production d'un discours conduirait à de graves déconvenues. On veut signifier qu'il n'y a pas de cas, de figure type, où telle méthode est indispensable, mais qu'au contraire la recherche de l'information à l'aide d'entretiens suppose qu'aient été faites préalablement des hypothèses de recherche sur un objet précisé.

Il n'est pas dans notre intention de faire un catalogue exhaustif de l'ensemble des techniques possibles car elles seraient fort nombreuses, mais non testées [1]. On a présenté ici quatre types de techniques qui ont fait leur preuve par ailleurs, dans une enquête ou une autre et qui, à ce titre, méritaient d'être citées.

Il faut également souligner que les techniques classiques et l'une ou l'autre des techniques particulières explicitées ne sont pas mutuellement exclusives. Il est en effet tout à fait possible que d'une part on décide d'utiliser la technique des « contre-exemples » et qu'au cours de l'entretien on se serve d'expressions brèves ou de relances de type « miroir » pour permettre à e. de progresser dans sa propre approche du problème.

1. On pense notamment à l'utilisation d'un discours sans signification fait par E., à l'utilisation de fantasmes, au décalage verbal, etc.

3.6. LE CHOIX, LA FORMATION ET LE CONTRÔLE DES ENQUÊTEURS

On aborde ici le domaine qui est à notre sens le plus hasardeux, au moins en ce qui concerne le choix. Ce qui va être dit quant aux modalités de choix des enquêteurs doit donc être considéré, dans la plupart des cas, comme indicatif et ne pas être pris comme une règle de recrutement ; nous serions en ce domaine (et en de nombreux autres aussi) incapables d'en donner.

3.6.1. Le choix des enquêteurs

On voudrait tout d'abord attirer l'attention sur le fait qu'il n'existe pas [1] d'interviewer né. Idée qui peut (et devrait) faire sourire lorsqu'elle est énoncée avec autant de naïveté, mais qui trouve son corrélat dans des expressions telles que « Oh ! moi, je sais faire parler les gens... c'est parce que les gens se confient à moi... c'est étrange, lorsque les gens sont en relation avec moi, ils se croient obligés de me raconter leur vie ». On conviendra que ces locutions sont relativement fréquentes dans le langage courant et recouvrent souvent une idéologie selon laquelle il y a des gens plus capables que d'autres de mener des entretiens, car ils auraient des qualités natives qui les prédisposent à cela.

Idéologie suffisamment fréquente pour produire des tentatives de validation expérimentale des caractéristiques propres à la « personnalité interviewante ». Dans une recherche américaine, on fait passer un TAT à 11 étudiants pratiquant l'entretien en néophytes, et à 14 interviewers professionnels. L'auteur trouve trois variables reliées positivement aux performances des différentes personnes : le sens des relations humaines, l'introspection et la réceptivité. Ce n'est ni le lieu ni l'objet, pour entamer ici un débat sur la science et sur ce qui est scientifique ou non ; on admettra cependant que les concepts de « sens des relations humaines » ou de « réceptivité » sont pour le moins mal définis, et qu'il apparaît difficile d'en faire l'amorce d'une théorie de « l'enquêteur né ». De plus, d'autres recherches montrent que s'il existe une relation positive entre des interviews réussies et la *personal sensitivity* (dont il faudrait encore démontrer qu'elle est identifiable à l'un des concepts exhumés ci-dessus) elle est pour le moins modeste.

En d'autres termes, il nous semble que le choix d'un enquêteur doit se faire à partir des deux critères : la compétence technique, et la situation constituée par une étude donnée sur une population donnée.

Pour ce qui est de la compétence technique, elle est aisément repérable à travers deux types de critères :

— La capacité de l'interviewer à mener une interview non directive [2], c'est-à-dire sa capacité à utiliser à bon escient les techniques nécessaires, à ne pas biaiser le discours du

---

1. A notre connaissance du moins, mais nous ne demandons qu'à apprendre.

2. On peut considérer que les interviews semi-directives et directives mettant en œuvre le même type de techniques (mais dans une situation moins complexe) ne sont que des cas particuliers, à ce niveau du moins, de l'entretien non directif et que, par conséquent, un enquêteur réalisant de façon satisfaisante les entretiens non directifs, pourra réaliser également les entretiens semi-directifs et les entretiens directifs.

sujet en y introduisant d'une façon ou d'une autre son propre cadre de référence, à obtenir un nombre d'informations suffisantes (ou estimées telles par le responsable de l'étude).

— L'adéquation entre le rôle d'enquêteur et l'ensemble des comportements verbaux et non verbaux de l'individu, c'est-à-dire en fait sa capacité à s'identifier à un professionnel de l'entretien et à disparaître en tant qu'individualité derrière ce rôle.

Il est tout à fait évident que ce dernier critère constitue à la limite un vœu pieux ou un idéal qu'il faut s'efforcer d'atteindre, mais il est garant d'un certain type de formation des enquêteurs. Cependant, parce qu'il est un idéal difficile à réaliser, il est nécessaire de tenir compte, dans la mesure du possible, d'une adéquation entre l'individualité de l'enquêteur et la spécificité de la situation.

Diverses études menées sur des caractéristiques physiques ou psychologiques ont montré qu'il y avait lieu de considérer avec soin la situation d'enquête pour éviter un certain nombre de mésaventures. Ainsi, les facteurs de race et de sexe ont été analysés et il a été montré qu'ils pouvaient être inhibiteurs ou faciliteurs suivant la situation. Un Blanc interrogeant des Noirs, aux Etats-Unis, sur les problèmes liés à la ségrégation, risque d'obtenir des entretiens moins riches qu'un Noir interrogeant des Noirs, un homme interrogeant des militantes du M.L.F. sur les vertus de la femme au foyer risque d'obtenir des réponses biaisées du seul fait qu'il est un homme, etc. Dans le même ordre de facteurs, on pourrait citer l'apparence physique, l'âge, le vêtement, etc.

Par ailleurs, les facteurs psychologiques tels que : le cadre de référence de l'enquêteur sur le sujet d'étude (existant ou qui se créera au fur et à mesure que le nombre des entretiens réalisés par lui croîtra) ; l'attitude de l'enquêteur vis-à-vis de la population qu'il va interroger ; le niveau culturel, la religion, doivent être connus et rapprochés de la situation d'enquête, etc.

Tout ceci conduit à conclure que :

— Les enquêteurs doivent être sélectionnés (rappelons que nous parlons ici *essentiellement* au niveau de l'entretien non directif et semi-directif, il n'en est pas de même pour un questionnaire standardisé) en fonction d'une étude précise et d'une population d'enquêtés particulière, afin que des caractéristiques diverses ne surdéterminent pas toute la production discursive de l'interviewé.

— Leur compétence technique doit être testée du double point de vue de l'utilisation qu'ils font des techniques d'entretien et de leur capacité à se confondre avec le rôle d'enquêteur, de devenir des professionnels de l'enquête et non, par exemple, « des étudiants qui font ça pour gagner leur vie, en attendant d'avoir une vraie profession ».

— Les qualités d'empathie, de sympathie, de sensibilité aux relations humaines, de réceptivité, etc., ne sont pas pertinentes dans l'état actuel des connaissances et que, de plus, quand bien même elles le seraient, ne sont pas nécessairement déterminantes pour la production d'un « bon » entretien (étant entendu qu'une attitude d'écoute attentive vaut mieux dans la plupart des cas que le désir forcené d'être sympathique à l'interviewé).

3.6.2. La formation des enquêteurs

Si l'on part de l'hypothèse que l'enquêteur n'est pas formé du tout, on peut tracer un programme de formation et présenter quelques instruments de formation.

— Les grandes têtes de chapitre (et notre intention n'est pas d'analyser davantage ce point) d'un tel programme pourraient être :

1. Un enseignement sur les entretiens (leurs caractéristiques, leurs biais, leur utilité, etc.) et de façon plus générale sur le recueil des données et leur traitement.
2. L'analyse des différentes techniques d'entretiens à travers un enseignement didactique et les exercices sur des enregistrements audio-visuels d'entretiens de différentes natures, sur différents sujets, avec différents enquêteurs. Exercice d'évaluation à la fois des comportements verbaux et des comportements non verbaux.
3. La pratique des entretiens tout d'abord en « laboratoire » avec enregistrement audio-visuel de l'entretien et évaluation de celui-ci, ensuite sur le « terrain » avec une phase d'évaluation particulièrement rigoureuse.
4. La prévision de formations complémentaires pour tel ou tel problème particulier nécessitant une technique particulière.

Un certain nombre d'instruments pédagogiques peuvent être utilisés durant cette formation ; on se contentera d'en citer trois :

Le jeu de rôle, consistant par exemple à croiser les rôles E.-e. pour une même personne qui sera tantôt interviewée, tantôt « interviewante ». Ce jeu de rôle ne doit pas se faire, à notre sens, en présence d'un auditoire, car on introduit, ce faisant, des variables non contrôlées telles que l'influence d'un auditoire attentif et silencieux sur le comportement des partenaires E.-e.. Il doit par ailleurs avoir lieu entre une personne et plusieurs partenaires dans des conditions différentes, définies par le jeu de rôle. Les enregistrements audio-visuels doivent être faits pour permettre *ensuite* une évaluation par l'ensemble des participants.

L'exercice ne consistant à utiliser qu'une partie des techniques d'entretien, ou encore l'exercice d'évaluation d'un cas (ici un enregistrement).

L'utilisation d'un « sujet » (une personne extérieure) réel, mais dans une situation de laboratoire, pour se prêter aux interviews. Ce sujet ayant été choisi pour permettre de mettre en évidence le maximum de biais possible.

3.6.3. Le contrôle des enquêteurs

Contrôler des enquêteurs suppose que l'on définisse au préalable ce que l'on entend par contrôler. S'il s'agit seulement de s'assurer que les enquêteurs font bien leur travail et mettre en place une personne pour ce faire, la notion de contrôle nous paraît restrictive.

En fait, le contrôle devrait porter sur l'« avant » et le « pendant ». « Avant », il s'agit de « contrôler » le cadre de référence de l'enquêteur quant au sujet qu'il est chargé de traiter. *Ceci suppose un entretien de l'enquêteur et son analyse*, afin de connaître ce cadre de référence, et également de le lui faire connaître pour qu'il se méfie de la teneur

de ses interactions au cours de la relation E.-e.. Ceci suppose également une connaissance de ses comportements en cours d'entretien, que l'entretien préalable peut fournir.

« Pendant », il faut s'assurer que, au fur et à mesure que le nombre des entretiens faits par une personne augmente, celle-ci ne constitue pas vis-à-vis du problème une série d'hypothèses ou un cadre de référence dans lequel elle fait entrer les enquêtés.

Les discours, dans ce cas, deviennent le reflet de la pensée de E.. Malheureusement ce phénomène est difficile à éviter autrement qu'en limitant le nombre d'entretiens faits par une même personne sur un sujet donné (de 5 à 8 entretiens par enquêteur).

Il faut contrôler également la qualité du travail, le respect de l'échantillonnage, etc., toutes choses connues sur lesquelles il n'est point besoin d'insister.

BIBLIOGRAPHIE

- Bingham W.V.D., Moore B.V., *How to interview*, New York, Harper and Row, 1924. Ouvrage « classique » qui marque plus ou moins les débuts des réflexions systématiques sur les méthodes d'entretien.
- Cannell C.F., Kahn M.L., *The dynamics of Interviewing, Theory, Technics and Cases*, New York, Wiley, 1957. Ouvrage plus récent sur les méthodes d'entretien.
- Cannell C.F., Kahn M.L., « Interviewing », *in Lindzey G., Aronson E., Handbook of social psychology*, Reading, Mass., Addison Wesley, 1968. Reprend pour partie l'ouvrage précédent, mais est fait davantage dans l'esprit « revue de questions », et actualise les données.
- Rogers C., *Counselling and psychotherapy*, Boston, Houghton Mifflin, 1942. Ouvrage « classique » pour ce qui concerne l'entretien clinique.

Chapitre 4

COMMENT INTERROGER ? LES QUESTIONNAIRES

4.1. APPROCHE QUALITATIVE ET APPROCHE QUANTITATIVE

Il est habituel de considérer qu'une enquête complète doit commencer par une phase qualitative, sous la forme d'un ensemble d'entretiens non directifs ou structurés, suivie d'une phase quantitative, l'application d'un questionnaire à un échantillon permettant une inférence statistique au cours de laquelle on vérifie les hypothèses élaborées au cours de la première phase et on les complète par des renseignements chiffrés. En effet, si seule une méthode tout à fait ouverte permet d'aborder un problème lorsqu'on n'a encore que peu d'hypothèses et qu'on sait mal comment ce problème se pose pour la population intéressée, en revanche, pour construire un questionnaire, il faut évidemment savoir de façon précise ce qu'on cherche, s'assurer que les questions ont un sens pour chacun, que tous les aspects de la question ont bien été abordés, etc. C'est cela qu'on tire des entretiens et du test des premières versions de questionnaire.

Mais on limiterait la portée de la phase qualitative si on n'en faisait qu'une étape de préparation. Malgré le petit nombre de personnes interrogées, certaines conclusions suffisamment solides peuvent en être tirées, notamment tout ce qui peut se ramener à un inventaire plus ou moins structuré d'attitudes, de représentations, de comportements, de motivations, de processus, etc. Même s'il n'est pas possible d'en inférer l'importance de chacune des modalités recensées, on peut, à partir d'entretiens, au moins

conclure à leur existence. D'autre part, certaines composantes affectives des situations étudiées sont souvent très difficiles à saisir par un questionnaire, faute d'un vocabulaire adéquat et adapté à tous les cas, et aussi parce que c'est une méthode qui permet facilement des rationalisations : une personne domine beaucoup mieux ses réponses à des questions précises que son discours face à un enquêteur non directif.

Sur ces points, on est donc souvent obligé, même si on juge que c'est insuffisant, de s'en tenir à la phase qualitative.

Les objectifs de chaque phase étant fixés, on peut s'apercevoir que la phase qualitative n'est pas nécessaire, parce qu'on dispose déjà de tout ce qu'elle pourrait apporter. C'est le cas lorsque le domaine a déjà été bien étudié, que le vocabulaire utilisé est connu, que les hypothèses ont déjà été élaborées, ou peuvent l'être à partir d'autres travaux.

Parfois, contrairement à l'ordre habituel, après une enquête quantitative, une phase qualitative complémentaire peut être utile, par exemple pour aider à interpréter certains résultats inattendus. On peut ainsi concevoir, au lieu de la succession classique, un va-et-vient entre des phases qualitatives et quantitatives, celles-ci servant alors à mettre en évidence des relations que celles-là permettront d'interpréter. Par exemple, dans le domaine des transports urbains, on a commencé par des enquêtes entièrement quantitatives qui ont mis en évidence les poids relatifs du temps et du prix comme déterminants du choix de mode de transport. Ce n'est qu'ensuite que des études par entretiens ouverts ont pu saisir la signification, pour les usagers, de ces variables et d'en mettre en évidence d'autres. Ce qui a permis de concevoir ensuite des enquêtes statistiques plus complexes.

Les objectifs d'un questionnaire peuvent se ramener à un petit nombre :

Estimer certaines grandeurs « absolues » : dépenses au cours d'une période donnée ; pourcentage de personnes ayant une certaine opinion ; pourcentage de personnes lisant un certain journal ou achetant un certain produit, etc.

Estimer des grandeurs « relatives » : par exemple lorsqu'on a élaboré une typologie, estimer la proportion de chaque type dans la population étudiée.

Décrire une population ou des sous-populations : par exemple donner les caractéristiques des acheteurs d'un produit, des lecteurs d'un journal, de ceux qui affirment une certaine opinion, etc.

Vérifier des hypothèses : sous la forme de relations entre deux ou plusieurs variables ; par exemple, vérifier si la nature ou la fréquence d'un comportement varie avec l'âge, si les opinions et les comportements relatifs à un certain objet sont cohérents, ou, souvent, des hypothèses plus complexes.

Ce dernier objectif est en général le plus important, pour les sociologues du moins, sinon pour les économistes. Il est rare, en effet, qu'un sociologue se pose les problèmes d'estimation de grandeurs. Cela tient principalement à la nature des variables avec lesquelles il travaille, qui n'ont que rarement un intérêt en elles-mêmes. Estimer la consommation d'un certain produit ou le nombre de chômeurs à un moment donné, comme le font les économistes, a un sens dans la mesure où l'on peut donner des définitions sans ambiguïté de ces termes ; même si des définitions concurrentes sont proposées, comme dans le cas du chômage, on sait que chacune correspond à une réalité diffé-

rente. En revanche, lorsqu'on étudie une opinion ou une attitude, on sait que, dans certaines limites, la distribution des réponses dépendra de la formulation des questions. Cela n'a donc guère de sens de chercher à estimer avec une grande précision la population de personnes répondant « oui » à une question, alors que la formulation de celle-ci comporte toujours une part inévitable d'arbitraire. En revanche, comparer la proportion des réponses « oui » des différents groupes sociaux, ou étudier l'évolution de cette proportion dans le temps est beaucoup plus justifié. On se trouve de ce fait renvoyé à un problème de mise en évidence de relations, donc de vérification d'hypothèses, plutôt que d'estimation.

Toutefois, à côté de ces raisons proprement techniques, l'absence d'intérêt de la majorité des sociologues, surtout en France, pour les estimations, provient également, dans bien des cas, d'un mépris, que certains croient scientifique, pour ce qu'ils appellent la « sociographie », c'est-à-dire la description, avec un faible souci de généralisation, des phénomènes sociaux. C'est une activité considérée comme inférieure, sans valeur théorique. Il y aurait beaucoup à dire sur les plans épistémologique et idéologique de cette distinction entre description d'une part, interprétation et théorie de l'autre, l'épithète prestigieuse de « scientifique » n'étant évidemment reconnue qu'à ces dernières. Outre que la distinction elle-même est souvent difficile à établir devant des exemples concrets, ce dédain et ce désintérêt pour la « description » traduisent, à notre sens, la permanence d'habitudes de pensée héritées des origines philosophiques de la sociologie : beaucoup de sociologues raisonnent comme si leur connaissance du détail du fonctionnement de leur société était déjà suffisante, et qu'il n'y ait plus qu'à réfléchir à partir de cette connaissance pour en tirer des lois générales ou, si l'on est prudent, des hypothèses de grande portée théorique. Mais si l'on pense qu'il y a encore à découvrir à ce niveau, que les « descriptions » qu'on peut donner actuellement ne sont pas suffisantes, alors l'intérêt des estimations, et d'ailleurs de toutes les formes de recherche empirique, redevient évident.

Réaliser les quatre objectifs que nous avons distingués suppose qu'on les ait clairement explicités, et qu'on ait rendu opérationnels les différents concepts utilisés, c'est-à-dire qu'on soit en mesure de faire correspondre à chaque concept une ou plusieurs réponses au questionnaire, réponses qui serviront de définitions opérationnelles ou d'indicateurs de ces concepts. Si l'on parle d'une attitude, par exemple, il faut préciser la règle permettant de classer chaque sujet dans une catégorie en fonction de ses réponses.

Contrairement aux différentes formes d'entretiens, la conception et la rédaction d'un questionnaire sont entièrement déterminées par l'exploitation statistique qui en est prévue [1]. Cela implique qu'on puisse valablement *dénombrer* les réponses à chaque question, c'est-à-dire qu'on puisse considérer effectivement comme équivalentes des réponses semblables. Cela semble une condition évidente, toujours satisfaite. Mais cela impose que ce soit le même questionnaire exactement qui soit posé à toutes les person-

1. Il ne faut toutefois pas prendre cette exigence trop au pied de la lettre : s'il est effectivement nécessaire d'avoir toujours à l'esprit la future exploitation statistique et ses contraintes, il faut aussi prendre soin de rendre le questionnaire acceptable, et même, si possible, agréable pour le sujet. Par exemple, on pourra lui poser une question ouverte qui lui permettra de s'expliquer librement, même si on pense que les réponses seront peu utilisables.

nes interrogées [1]. Une fois que le travail sur le terrain est commencé, il est donc exclu d'apporter des modifications quelconques à l'énoncé des questions ou à leur ordre, même si l'on est convaincu qu'il s'agit d'améliorations importantes, même si l'on s'est rendu compte d'erreurs graves. Dans ce dernier cas, il faudrait, dans l'idéal, recommencer entièrement l'enquête avec le questionnaire corrigé, et considérer les premiers sujets interrogés comme nuls. Si l'on ne dispose pas de suffisamment de moyens ou de temps pour cela, on poursuivra avec une version corrigée, mais en sachant que certaines questions n'avaient été posées qu'à une partie de la population, et que tous les dépouillements dans lesquels elles devaient intervenir ne pourraient se faire que sur ce demi-groupe de sujets. Ce qu'il faut absolument éviter, c'est de confondre les deux formes du questionnaire.

C'est pour la même raison de maintien de la constance des conditions de passation qu'il ne faut pas que l'enquêteur se trouve obligé d'expliquer certaines questions à une partie des sujets. Le questionnaire doit être conçu de telle sorte qu'il n'y ait aucun besoin d'explications autres que celles qui sont explicitement prévues.

La construction du questionnaire et la formulation des questions constituent donc une phase cruciale du déroulement d'une enquête. On ne peut pas laisser certains points dans le vague, en se disant qu'on précisera plus tard, au vu des réponses. *Toute erreur, toute maladresse, toute ambiguïté, se répercuteront sur l'ensemble des opérations ultérieures, jusqu'aux conclusions finales.*

Dans le cas des entretiens, le problème de maintien des conditions constantes se pose tout autrement. Certes, et nous avons déjà insisté sur ce point, la standardisation de la consigne et des modes d'intervention de l'enquêteur est nécessaire. Mais la notion même de comparabilité entre sujets prend un sens plus vague que dans les questionnaires, puisque chacun suit en principe sa propre ligne de pensée, ce qui fait que des entretiens menés avec la même méthode et à partir de la même consigne peuvent se dérouler de façon complètement différente. La comparaison n'est plus possible point par point, mais seulement globalement. D'autre part, au cours de l'analyse et de l'interprétation, chaque entretien est considéré en lui-même, et dans sa totalité ; on peut donc toujours, en principe, tenir compte par exemple d'un éventuel changement de consigne, ou de réactions propres à un sujet particulier. La systématisation stricte, tout en étant souhaitable, n'est pas une condition aussi absolue que dans le cas des questionnaires.

Une autre conséquence du caractère statistique de l'exploitation d'un questionnaire c'est que le chercheur n'a jamais connaissance de l'ensemble des réponses d'un même individu ; même s'il a la curiosité de lire intégralement quelques questionnaires, ce qui est très « parlant », ils resteront en très petit nombre et il n'en tirera rien de plus que quelques impressions. La matière sur laquelle il travaillera réellement consistera en tableaux statistiques donnant les croisements, dans l'ensemble de l'échantillon ou dans certains sous-échantillons, en réponses à deux, trois ou quatre questions, en pratique

1. Cette condition n'est pas toujours remplie. Par exemple, certaines questions n'ont de sens que pour une partie de l'échantillon (des questions sur leur mariage, par exemple, ne peuvent être posées qu'à des personnes mariées, veuves ou divorcées). Ou encore, lorsqu'on construit des questionnaires « en arbre » certaines questions sont posées en fonction des réponses aux questions précédentes. Pour chacun de ces cas, souvent inévitables, il faut se demander si ces variantes ne risquent pas d'affecter les réponses aux questions communes suivantes, et donc la comparaison entre les groupes.

jamais plus. Cela implique qu'il n'est pas possible en général de vérifier des hypothèses faisant intervenir un plus grand nombre de variables, ni de compter sur des recoupements pour corriger des erreurs ou expliquer les réponses étonnantes.

Cette limite n'a en soi rien de fondamental, elle est la conséquence de deux difficultés pratiques. Premièrement l'interprétation des résultats devient très complexe lorsque le nombre de variables prises en considération croît, et le plus souvent on se trouve démuni pour interpréter des tableaux croisant un trop grand nombre de questions, sauf dans les cas, encore très rares dans la pratique des enquêtes, où l'on disposerait à l'avance d'hypothèses précises, formalisées en un modèle (par exemple, un modèle de causalité) qu'on pourra vérifier globalement. Mais si les résultats amènent à rejeter le modèle, on se retrouve dans une situation où l'information est trop complexe pour être traitée par des interprétations informelles.

La deuxième limite tient au nombre de sujets nécessaires. Celui-ci croît de façon exponentielle avec le nombre de questions prises en compte simultanément ; on arrive donc très vite à des échantillons énormes, pratiquement impossibles à obtenir.

Des méthodes récentes, dites d'analyse de données, en général dérivées de l'analyse factorielle et fondées sur un usage intensif des possibilités de l'informatique, permettent de résumer les réponses à un grand nombre de questions considérées simultanément. Mais les postulats de ces méthodes sont en général très forts, et surtout ils constituent plus des méthodes de description systématique et de mise en ordre de données complexes que de vérifications d'hypothèses, qui ne sont en général pas formulées de façon qui les rende comparables à ce qu'on obtient par ces méthodes. Sans du tout nier leur intérêt, il faut reconnaître que leur champ d'application reste encore très limité.

Lorsqu'on travaille sur un questionnaire, il n'y a pas, comme lorsqu'on travaille sur des entretiens libres ou structurés, la saisie plus ou moins systématique ou intuitive de la cohérence d'un individu seul. Dans ce cas, nous le verrons (cf. 6.6) le problème est de passer d'une collection d'individus à la population, d'agréger un ensemble de discours portant sur des individus en un discours unique portant sur la population. Dans le cas des questionnaires, le problème est inverse : l'agrégation est automatique, si l'on peut dire, puisqu'on travaille toujours sur la distribution et qu'on s'est assuré, par la standardisation du recueil des réponses, que la constitution de ces distributions était légitime. Mais, ce faisant, on perd la spécificité de chaque individu, on l'émiette en un grand nombre de grappes de réponses, ce qui peut amener à des erreurs d'interprétation sur lesquelles nous reviendrons.

Il faut toutefois éviter ici un malentendu. Si nous parlons de conserver la spécificité de chaque individu, il ne s'agit pas de considérer chacun comme absolument unique et incomparable aux autres. Tout effort scientifique vise précisément, sans se prononcer sur cette unicité, à énoncer des conclusions de portée générale. Le danger contre lequel nous mettons en garde ici, c'est celui de la séparation des réponses, qu'on isole des autres réponses données pour le même sujet. Mais il est évident que si chaque individu est considéré comme unique, si l'on ne peut pas trouver de point sur lequel il est comparable aux autres, tout discours général devient impossible, et avec lui toute science de l'homme.

4.2. LES GRANDES LIGNES D'UN QUESTIONNAIRE

Un questionnaire, par définition, est un instrument rigoureusement standardisé, à la fois dans le texte des questions et dans leur ordre. Toujours pour assurer la comparabilité des réponses de tous les sujets, il est absolument indispensable que chaque question soit posée à chaque sujet de la même façon, sans adaptation, ni explications complémentaires, laissées à l'initiative de l'enquêteur. Pour que cela soit possible, il faut évidemment que la question soit parfaitement claire, sans aucune ambiguïté, et que le sujet sache exactement ce qu'on attend de lui. Si, pour éviter des malentendus, ou des réponses manifestement absurdes, les enquêteurs sont obligés de reformuler une question pour certaines personnes, ou l'expliquer, c'est que la question est mauvaise.

Sans tomber dans l'excès de ceux qui prétendent que, par une formulation habile des questions, on peut obtenir n'importe quelle conclusion d'une enquête, il faut être conscient du fait que le libellé de la question affecte certainement la réponse, dans des limites qui sont d'ailleurs variables selon le contenu. Une bonne question ne doit donc suggérer aucune réponse particulière, ne doit exprimer aucune attente, ne doit rien exclure de ce qui peut venir à l'esprit à qui on la pose.

Malgré son importance, la rédaction du questionnaire reste dépendante du savoir-faire et de l'expérience du chercheur. Il n'est pas possible, actuellement, d'énoncer des règles pour la construction du questionnaire et la manière de rédiger les questions. Au mieux on peut énumérer un certain nombre de mises en garde, donner une liste de points auxquels il est bon d'avoir pensé. Mais ces conseils restent le plus souvent négatifs, et ils ne découlent que rarement d'études empiriques systématiques qui assureraient leur validité. Ce sont plutôt des conseils pratiques qui se transmettent dans la profession, appuyés sur le « bon sens » et l'expérience. Dans quelques cas on a tenté expérimentalement de comparer diverses formulations, mais les conclusions qu'on a pu en tirer, malgré leur intérêt, sont encore loin de couvrir tout le champ, et d'apporter des réponses à tous les problèmes que la pratique peut poser.

A côté du texte des questions, l'ordre dans lequel on les pose est également très important. Arrivé à un certain point du questionnaire, les questions qui précèdent ont donné au sujet, qu'on le veuille ou non, une idée du champ couvert par l'enquête, l'ont familiarisé avec le thème et la façon particulière dont il est abordé, lui ont donné l'occasion d'y réfléchir. On lui aura fait voir des aspects du problème auxquels il n'avait peut-être pas pensé. On lui aura très probablement demandé de donner son avis sur des opinions très différentes de la sienne propre : même s'il avait toutes les possibilités d'exprimer son désaccord, son cadre de référence a pu en être modifié. Une même question, placée au début ou à la fin d'un questionnaire, avant ou après une certaine autre question, pourra donc susciter des réponses différentes.

Au cours d'une enquête menée dans le cadre d'une étude de motivation sur les caméras d'amateur, on a, par inadvertance, posé la même question tout au début du questionnaire, et à la fin. C'était la question « envisagez-vous au cours des prochaines années d'acheter une caméra ? » (question posée à ceux qui n'en possédaient pas). A la fin d'un questionnaire de 30 minutes environ, le nombre des réponses positives à cette

même question avait presque doublé. C'est que, entre temps, on les avait fait parler de vacances, envisager des utilisations possibles de la caméra, etc. ; ce qui n'était qu'une possibilité lointaine au début, était devenu familier après une demi-heure. Chaque question doit être interprétée en référence à sa position dans le questionnaire, et donc cette position doit être soigneusement choisie.

Un questionnaire doit apparaître comme un échange verbal, aussi naturel que possible. Les questions doivent s'enchaîner l'une à l'autre sans redites et sans coq-à-l'âne. Ceux-ci toutefois sont souvent difficiles à éviter, par exemple lorsque les hypothèses élaborées pour rendre compte d'un phénomène font appel à des notions qui, aux yeux de la plupart des sujets, n'ont rien à voir avec lui. Par exemple, on peut chercher à expliquer certains aspects de la consommation par les caractéristiques de l'entreprise où l'on travaille. Le sujet à qui l'on a annoncé une enquête sur la consommation, qui aura répondu à toute une série de questions sur ce thème, pourra être dérouté si, sans transition, on l'interroge sur son travail. Il est bon de prévoir, à ce moment, une phrase d'avertissement, indiquant, lorsque c'est possible, pourquoi on va poser des questions inattendues, ou du moins annonçant qu'on va changer de thème.

Lorsqu'on effectue des enquêtes par quota, il est nécessaire de commencer par les questions permettant de savoir si la personne qu'on a en face de soi entre ou non dans les catégories voulues. C'est indispensable, mais cela ne constitue pas pour autant une bonne entrée en matière. Ces questions sont jugées trop personnelles, et on ne voit pas pourquoi on les pose alors que, à la fin du questionnaire lorsque la relation entre l'enquêteur et l'enquêté est bien établie, que ce dernier a déjà répondu à des questions très diverses, elles sont beaucoup mieux acceptées. Si commencer ainsi est inévitable, une phrase introductive du genre « avant de commencer, je dois vous poser quelques questions sur vous-même » peut améliorer le contact.

Les premières questions sont très importantes. Ce sont elles qui indiquent aux personnes enquêtées le style général du questionnaire, le genre de réponse qu'on attend d'elles, le thème qui va être abordé. C'est aussi sur elles que s'établit la relation enquêteur-enquêté, que le sujet aura l'impression qu'on empiète ou non dans sa vie privée ; il est souvent préférable de commencer par des questions susceptibles d'intéresser le sujet, sans l'effrayer.

En commençant par une question fermée, on risque de donner l'impression qu'on n'attend que des réponses brèves, sans détails, et qu'on ne s'intéresse qu'à ce qui entre dans des cadres pré-établis. Ce n'est pas grave si le questionnaire est constitué uniquement par des questions fermées mais, si ce n'est pas le cas, une question ouverte arrivant pour la première fois après une longue suite de questions fermées ne suscitera que des réponses brèves et pauvres.

En principe, il y a avantage à regrouper toutes les questions portant explicitement sur le même thème. Sinon le sujet pourra avoir l'impression qu'on ressasse, qu'on lui demande quelque chose qu'il a déjà dit, qu'on tourne en rond. Toutefois, lorsqu'il s'agit de questions d'opinions, de préférences, d'attitudes, etc., il est parfois nécessaire de passer outre à cette recommandation pour éviter que les réponses ne soient biaisées par le souci de cohérence.

S'il faut assurer une certaine cohérence dans le contenu des questions et dans la succession des thèmes, la variété dans la *forme* des questions est, en général, bien accueillie, elle évite l'impression de monotonie qui constitue l'un des principaux dangers des questionnaires qui se prolongent.

La durée acceptable d'un questionnaire dépend beaucoup de l'intérêt que le sujet porte au thème et à la manière dont il est abordé, et aussi des conditions de passation ; la tendance des chercheurs est, en général, de vouloir poser le plus de questions possible, parfois sans penser qu'il faudra non seulement faire passer le questionnaire, mais aussi l'exploiter. Devant chaque question qui vient à l'esprit il faudra se demander de façon très précise ce qu'on en fera. Préparer à l'avance la phase de dépouillement permet souvent de se rendre compte que certaines questions, apparemment intéressantes, seront inutiles, soit qu'elles fassent à peu près double emploi avec d'autres, soit qu'elles soient équivoques en elles-mêmes et qu'elles exigeraient plusieurs autres questions complémentaires pour être interprétables.

Un questionnaire composé en majorité de questions fermées, ne devrait pas dépasser 45 minutes lorsque la passation se fait dans de bonnes conditions, c'est-à-dire au domicile du sujet ou dans un lieu tranquille. Au-delà, l'intérêt faiblit, ce qui se remarque à des signes comme la brièveté des réponses aux questions ouvertes et à la rapidité des réponses, qui indique que la réflexion sur la réponse a été faible. Toutefois, si le thème tient le sujet à cœur, le questionnaire peut être plus long, mais ne devrait sûrement pas dépasser l'heure.

Lorsque l'enquête se fait par échantillonnage sur place, dans la rue, ou dans des lieux publics, l'enquêté et l'enquêteur restant debout, avec autour d'eux des personnes qui parlent ou qui risquent d'intervenir, dix minutes d'interrogation constituent un maximum.

Certaines questions sont particulièrement ennuyeuses. C'est par exemple le cas de celles qui exigent des descriptions minutieuses de comportements peu intéressants en eux-mêmes, ou à propos desquels les sujets voudraient dire autre chose que ce qu'on leur demande. En revanche, les questions sur les opinions et les jugements de valeur sont, en général, bien accueillies. Les personnes interrogées ont l'impression qu'on leur demande vraiment leur avis, et cela correspond généralement bien à l'idée qu'elles se font des enquêtes.

Les demandes d'explication d'une réponse, sous la forme de « pourquoi » ajoutés à chaque question, suscitent facilement l'irritation de personnes qui ont beaucoup de peine à répondre, et qui jugent souvent de telles questions indiscrètes, surtout si elles se répètent.

Si, comme cela arrive parfois dans une recherche complexe, le nombre de questions nécessaires, même quand on les a examinées d'un œil critique, reste trop élevé, il reste la possibilité de les partager en deux ou trois groupes, et d'en faire plusieurs questionnaires distincts qui seront appliqués à des échantillons comparables mais indépendants. Le défaut le plus évident de cette manière de faire, c'est qu'elle double ou triple le nombre de sujets nécessaires, donc le coût de l'enquête. Mais, et c'est parfois plus grave, certains dépouillements deviennent impossibles : on ne peut pas croiser deux questions qui n'ont pas été posées aux mêmes sujets. Avant de se résoudre à dédoubler un ques-

tionnaire, il est donc indispensable de vérifier que toutes les exploitations prévues sont possibles. Lorsque cette exploitation prévoit qu'un petit nombre de variables explicatives seront croisées avec toutes les autres questions et qu'il n'y aura pas de croisements entre celles-ci, il suffira de poser les questions relatives aux variables explicatives dans les deux questionnaires.

Une solution intermédiaire, qui évite ces inconvénients, consiste à voir la même personne deux ou plusieurs fois. Le danger, dans ce cas, réside dans les aspects temporels qu'on maîtrise très mal : on ne sait pas bien ce qui peut intervenir entre la passation des deux questionnaires. Certaines personnes peuvent réfléchir au thème, d'autres n'y plus penser. Une maturation peut se produire, qui modifiera le point de vue du sujet, ou des événements oubliés peuvent revenir à l'esprit. Des précautions sont donc nécessaires.

4.3. LES DIFFÉRENTS TYPES DE QUESTIONS

En ce qui concerne leur *contenu*, on peut distinguer deux grandes catégories de questions :

— Celles qui portent sur des *faits*, en principe susceptibles d'être connus autrement que par enquête. Par exemple, on pourra demander à la personne interrogée quel journal elle a lu la veille, à quel endroit elle travaille, où elle a passé ses dernières vacances, etc.

— Celles qui portent sur des *opinions, des attitudes, des motivations, des préférences*, etc. Ce pourra être « estimez-vous que ce journal est objectif ? », « Pourquoi avez-vous passé vos vacances à tel endroit ? », « Le trajet jusqu'à votre travail est-il fatigant ? », etc. Cet ensemble très hétéroclite de questions qui ont en commun de porter sur des points qu'il est impossible de saisir autrement, est souvent appelé question « psychologique » ou « subjective ». Ces termes sont équivoques : un comportement est aussi « psychologique » qu'une opinion, et toute enquête essaie d'étudier des attitudes « objectivement ». Faute de mieux, nous les appelons *questions d'opinion*, en sachant que ce terme recouvrira des questions très diverses.

En plus de cette classification selon le contenu, les questions peuvent aussi se distinguer par leur forme. On opposera ainsi :

— Les questions *ouvertes*, auxquelles le sujet répond comme il le désire, donne les détails et les commentaires qu'il juge bons, et utilise son propre vocabulaire, ce qu'il dit étant en général noté intégralement par l'enquêteur.

— Les questions *fermées*, où l'on présente au sujet, après lui avoir posé la question, une liste pré-établie de réponses possibles, parmi lesquelles on lui demande d'indiquer celle qui correspond le mieux à celle qu'il souhaite donner.

Enfin, un questionnaire peut comporter autre chose que des questions au sens strict. On pourra demander aux sujets d'ordonner des objets [1], de leur attribuer des notes, de constituer des catégories, d'effectuer des mises en correspondance, etc. Des techniques

1. Le terme « objet » est évidemment pris ici dans son sens le plus large. Ce pourra être aussi bien des personnes, des opinions, des situations, etc.

très générales, comme les comparaisons par paires, ou plus spécifiques, comme le différentiel sémantique, voire des tests projectifs, sont aussi utilisées dans le cadre de questionnaires.

Toutefois, dans ce qui suit, lorsque nous ne le préciserons pas, c'est toujours des questions proprement dites qu'il s'agira.

4.4. QUESTIONS OUVERTES ET QUESTIONS FERMÉES

Par définition, dans un questionnaire, toutes les questions sont formulées à l'avance et l'enquêteur doit les poser sans adaptation ni explication. C'est à cette condition que les réponses des différents sujets peuvent être considérées comme comparables, qu'on peut les dénombrer et qu'elles se prêtent à une exploitation statistique. On peut pousser la standardisation plus loin, et uniformiser également les réponses en demandant à chaque sujet de choisir la sienne dans une liste pré-établie. Dans ce cas, la question est dite *fermée*. Sinon, si le sujet donne librement sa réponse, elle est dite *ouverte*. La plupart des questionnaires contiennent, en proportion variable, les deux formes de questions.

En pratique, lorsque la question est fermée, on présente au sujet la liste des réponses prévues inscrites sur une fiche, en précisant ce qu'on attend de lui. La consigne peut prendre des formes très différentes. Les plus usuelles sont :

— Indiquer *la réponse* la plus adéquate.
— Indiquer *plusieurs réponses* adéquates, le nombre de celles-ci étant *libre*.
— Indiquer *plusieurs réponses*, le nombre de celles-ci étant fixé.
— *Ordonner toutes les réponses*, de la plus à la moins adéquate.
— *Ordonner les n* (nombre fixé) *réponses* les plus adéquates.

Selon le cas, l'information recueillie sera plus ou moins riche, et la tâche demandée au sujet sera plus ou moins difficile.

La liste des réponses peut se réduire à la simple alternative oui/non, ou être très longue. Il est bon de prévoir dans cette liste deux rubriques supplémentaires : « ne sait pas » et « autres réponses », à quoi il faudrait ajouter parfois le refus de la question telle qu'elle est posée, ou le simple refus de répondre (cf.7.1.4.).

Au moment de l'exploitation, il sera important de prêter attention aux différentes réponses imprévues. En trop grand nombre, elles sont le signe que la liste des réponses a été mal conçue, ou même que la question a été mal formulée. Toutefois, les personnes interrogées sont en général très dociles, trop dociles, et elles acceptent facilement de se placer dans le cadre qui leur est proposé, même s'il ne correspond pas exactement à ce qu'elles pensent.

Une question ouverte soulève, à première vue, moins de difficultés. On la pose simplement, et on enregistre textuellement la réponse. Mais ensuite il faut coder ces réponses, c'est-à-dire les regrouper dans un petit nombre de catégories, ce qui pose parfois des problèmes délicats d'analyse de contenu.

Du point de vue de l'analyse des résultats, les questions fermées sont a priori les plus commodes. Tous les sujets ayant eu sous les yeux la même liste de réponses, ils sont immédiatement comparables. Lorsqu'il s'agit d'une enquête de passation et d'exploita-

tion rapides, comme un sondage d'opinion, on s'efforce de n'avoir que ce type de question. Dans les enquêtes plus approfondies, le choix se pose pour chaque question. Les raisons de choisir une forme ou l'autre sont nombreuses.

La raison la plus fréquente, mais pas nécessairement la meilleure de laisser une question ouverte, c'est qu'on n'a pas pu élaborer la liste des réponses à proposer aux sujets, soit parce que les pré-tests du questionnaire étaient insuffisants, soit parce que les réponses obtenues à ces pré-tests paraissaient trop complexes pour pouvoir être résumées en une liste de longueur acceptable. On peut avoir ainsi l'illusion de respecter la richesse de la pensée des personnes interrogées, et de recueillir un matériel plus complexe que celui que l'on recueillerait si la question était fermée. Mais il ne faut pas oublier qu'il faudra bien, au moment de l'exploitation, accepter une simplification et ramener, par le codage, toute la diversité des réponses à un petit nombre de catégories. Ce n'est donc que reculer devant la difficulté, sauf si on prévoit, en plus du codage conçu en vue de l'analyse statistique, une autre analyse plus qualitative. Toutefois, lorsqu'on a prévu de fermer une question et qu'on ne réussit pas à élaborer une liste de réponses acceptables, c'est souvent le signe qu'elle est mal posée.

D'autres raisons également peuvent, avec plus de justifications, amener à préférer laisser une question ouverte. La première est qu'un questionnaire entièrement fermé, surtout s'il est long, devient vite fastidieux. Soutenus par les listes de réponses qu'on leur présente les sujets peuvent réfléchir de moins en moins et prendre de moins en moins garde à ce qu'ils disent. Certains peuvent aussi s'irriter de voir qu'on ne les laisse pas s'exprimer librement, à leur manière, et avoir l'impression qu'on les manipule en leur imposant de rester dans un cadre qu'ils jugent inadéquat [1].

Introduire quelques questions ouvertes donnera au sujet l'impression, justifiée ou non, qu'on l'écoute vraiment. De plus, ces questions peuvent servir de garde-fou pour celui qui interprète les résultats, s'il veut bien se donner la peine de lire la transcription textuelle d'au moins une partie des réponses. Il ne faut toutefois pas trop compter sur cette possibilité d'élargissement : les réponses à une question ouverte succédant à une longue série de questions fermées sont souvent très pauvres, les sujets ayant pris l'habitude de donner des réponses brèves d'une part, et d'ordonner ou de choisir dans un matériel tout prêt d'autre part. Cette dernière conduite a souvent plus d'importance qu'on ne le pense dans la pauvreté des réponses aux questions ouvertes. Si celles-ci ne sont qu'un prétexte, cela n'est pas grave ; si par contre on veut les utiliser, le dosage et la succession des questions ouvertes et fermées prennent une grande importance.

Une autre raison de choisir la forme ouverte, c'est qu'elle laisse la possibilité de plusieurs codages différents, l'un portant par exemple sur le contenu de la réponse, ou sur les aspects de ce contenu, un second sur un autre aspect ou sur la forme, etc. Par exemple, dans une enquête sur la connaissance de la signalisation routière, on présentait chaque panneau au sujet en lui demandant d'en indiquer la signification. Les réponses ont été codées d'une part selon la gravité de l'erreur du point de vue de la sécurité, ce codage étant le même pour tous les panneaux, et d'autre part, selon le type particulier d'erreur commise, les catégories utilisées étant alors spécifiques à chaque cas.

1. Impression souvent parfaitement légitime. Nous verrons plus loin que le choix des réponses proposées peut influencer fortement la réponse donnée.

En choisissant entre les deux formes de questions, il faut tenir compte du fait qu'une liste de réponses fournit, qu'on le veuille ou non, un cadre de référence, et rappelle au sujet des aspects du problème auxquels il n'avait peut-être pas pensé spontanément. Selon le cas, ce peut être considéré comme un avantage ou un inconvénient. Parfois, on peut craindre qu'un sujet, devant une question ouverte, n'oublie certains aspects et ne donne la première réponse qui lui vient à l'esprit, sans que ce « choix » puisse être considéré comme significatif. Une liste de réponses couvrant la totalité du champ permet d'éviter ce biais. Dans d'autres cas, au contraire, on peut craindre que certaines réponses auxquelles le sujet n'aurait pas pensé spontanément lui semblent sur le moment séduisantes, et qu'il les choisisse sans beaucoup de réflexion.

En proposant une liste on fournit au sujet des indications sur le champ des réponses qu'on juge acceptables. Ce peut être un moyen de faire apparaître des réponses que, autrement, certains sujets n'auraient pas osé donner, par pudeur, par crainte d'apparaître comme non conformes, etc. A une question ouverte, seule une proportion minime de personnes donnera une réponse dont le contenu sexuel par exemple est manifeste. Cela ne suffit pas pour surmonter toutes les réactions de crainte mais élargit néanmoins le champ des réponses obtenues et aussi, parfois, l'oriente. Sans aborder des thèmes aussi délicats, signalons comme exemple une question posée sur les avantages de la voiture. Sous forme ouverte on n'obtient que des réponses très vagues, ou alors parfaitement rationnelles. En proposant une liste où, à côté de ces avantages rationnels, en apparaissent d'autres moins « avouables », comme le plaisir de la vitesse, on s'aperçoit qu'une fraction non négligeable des sujets le désignent, surtout si la consigne leur permet de donner plusieurs réponses. Cette fraction reste encore faible par rapport à celle qu'on peut supposer à partir d'entretiens non directifs, mais elle est suffisante pour confirmer au moins que ce plaisir peut être reconnu, et pour qu'on puisse en déduire les caractéristiques de ceux qui acceptent de dire qu'ils recherchent ce plaisir.

Si une liste pré-établie de réponses peut ainsi étendre aux yeux du sujet le champ de l'acceptable, il peut aussi évidemment le restreindre, et c'est là une source de biais très fréquente. Nous reviendrons en détail plus loin (cf. 4.5.3.3.) sur la formulation de ces listes de réponses, en passant en revue quelques-uns des risques encourus quant à l'influence exercée de cette manière sur les sujets. Si l'on sait bien que la formulation d'une question peut affecter fortement les réponses obtenues, le choix des réponses proposées est tout aussi important. Pour éviter ce risque on ajoute en général à la fin de la liste une rubrique « autres réponses » en demandant à l'enquêteur de transcrire intégralement cette réponse non prévue. Mais il ne faut pas trop compter sur ce correctif : l'expérience montre que très peu de sujets l'utilisent. Une réponse déjà formulée est plus attirante que l'idée encore confuse qu'on peut avoir.

Lorsqu'il s'agit de questions factuelles, portant sur les points ne posant pas de problèmes, utiliser des questions fermées présente l'avantage d'obtenir que tous les sujets utilisent la même nomenclature, et le même degré de finesse et de détail dans leurs descriptions. Par exemple dans une enquête sur les budgets de ménages, on pourrait demander simplement à chaque personne interrogée de noter quotidiennement ses dépenses, en lui laissant le choix des termes utilisés. Les réponses obtenues pourraient alors

être très différentes les unes des autres. Certaines seront trop détaillées pour les besoins de l'enquête, et donc demanderont un travail de codage supplémentaire, d'autres ne le seront pas assez, et risquent d'être inutilisables. Un même achat pourra être indiqué par une personne comme « légumes » et par une autre comme « petits pois très fins en conserve de la marque X », alors que l'information voulue par l'enquêteur était peut-être « conserve de légumes ».

Une liste donnée à la personne enquêtée évitera ces problèmes, et présentera l'avantage de servir de rappel. En effet, on peut penser que, si elle est suffisamment complète, elle permettra d'éviter certains oublis.

Devant les difficultés du choix entre ouvrir ou fermer une question, on recourt parfois à la formule suivante. On pourra, par exemple, poser la question sous forme ouverte et donner à l'enquêteur une liste de catégories dans lesquelles il codera immédiatement la réponse obtenue. Cela revient, en fait, à éviter une opération de codage distincte, l'enquêteur, étant sur place et ayant tous les indices supplémentaires fournis par le sujet (ton de la voix, hésitation, etc.), pourra éviter certaines erreurs. Mais on perd évidemment la possibilité, par exemple, d'effectuer plusieurs codages différents, qui constitueraient une tâche trop longue à effectuer sur place. On renonce également à la sécurité du double codage par deux personnes différentes.

Lorsque nous traiterons de l'analyse statistique des réponses, nous nous référerons naturellement aux réponses à des questions fermées. Pour que les réponses aux questions ouvertes puissent être exploitées de la même manière, il faut les coder, c'est-à-dire regrouper les réponses, toutes différentes, qu'on aura recueillies, en un petit nombre de catégories qui seront ensuite traitées de la même façon que les réponses aux questions fermées.

Pour qu'un code correct, avec une perte d'information minimale, soit possible, il faut pouvoir disposer du texte exact des réponses. Il faut donc que l'enquêteur note ce qui lui est dit, sans effectuer de tri, de résumé, ou de simplification. En effet, si l'on a laissé la question ouverte, il est possible que ce soit à cause des complexités du codage, qu'on était incapable d'effectuer à l'avance. On ne sait donc pas toujours quels seront les traits pertinents qu'on retiendra pour coder chaque réponse. Si l'enquêteur, en notant la réponse, fait lui-même un tri, ne retenant que ce qui lui semble important, il risque de laisser échapper des détails dont l'intérêt ne se révélera que plus tard. Par exemple, il peut estimer négligeable le fait qu'un sujet manifeste ou non une hésitation, un doute, et n'enregistrer que la réponse elle-même, ou plutôt son « contenu ». Dans bien des cas, cette hésitation n'aura en effet aucune importance, mais on ne peut pas exclure a priori que des codages portent précisément sur ce point, et qu'il se révèle utile, pour l'interprétation des résultats, de distinguer les personnes sûres de leur réponse et les autres. Il faut alors être en mesure d'effectuer la distinction.

A ces subtilités, on peut objecter que le sujet lui-même « code » sa réponse, qu'il ne dit certainement pas tout ce qui lui vient à l'esprit, particulièrement si les questions qu'on lui pose sont brèves et précises, et surtout que chacun ne « code » pas de la même manière. C'est certainement un point très important dont il est nécessaire de tenir compte dans l'interprétation des résultats. Il ne faut pas en conclure que le codage est impossible, mais se demander si les traits pertinents qu'on retient ont des chances

d'avoir été exprimés par chacun, et si on est capable de les interpréter correctement. On peut aussi donner à chaque sujet des indications suffisamment claires sur ce qu'on attend de lui afin, dans la mesure du possible, d'uniformiser ce « codage » par lui-même.

Coder les réponses à une question ouverte renvoie donc à un problème d'analyse de contenu, aussi complexe que l'analyse des entretiens non directifs. Les principales différences résident dans le fait que le matériel à analyser est très limité et que les contraintes de l'exploitation statistique exigent qu'on ramène la diversité des réponses individuelles à un petit nombre de catégories. Le moment crucial se trouve dans la construction du code, c'est-à-dire lors de l'établissement de la liste des catégories avec leurs définitions.

Une des indications de la qualité d'un code, outre évidemment son adéquation au problème étudié, c'est sa facilité d'emploi. S'il faut recourir trop souvent à des approximations ou à des interprétations pour décider dans quelle catégorie placer chaque réponse, si l'on hésite trop souvent entre plusieurs catégories, s'il reste beaucoup de réponses inutilisables, c'est que le code n'est pas bon.

Si un très petit nombre de catégories regroupe un très grand nombre de sujets, et d'autres très peu, le code sera peu discriminatif et il faudra chercher où le modifier en introduisant des distinctions dans les catégories trop nombreuses et en regroupant les autres. Si ce n'est pas possible, c'est que la question était mal posée, ou qu'elle portait sur un point sur lequel la population était trop homogène pour qu'on puisse l'utiliser.

4.5. LA FORMULATION DES QUESTIONS

4.5.1. Problèmes généraux

Une question a pour but d'acquérir une certaine information. Il faut donc la poser de manière qu'on soit certain que c'est bien cette information qu'on obtiendra, et elle seule. Il n'existe pas, pour cela, de règle générale, de recettes qu'on n'aurait qu'à appliquer ; tout ce qu'il est possible de faire, c'est de signaler les points auxquels il faut prendre garde, certaines erreurs à éviter. Mais cela n'exclut pas la réflexion, et on peut toujours imaginer des situations où un chercheur, pour répondre aux objectifs précis de son enquête, utilisera une formulation habituellement déconseillée.

Lors de la conception d'un questionnaire, d'ailleurs, il importe de conserver une attitude empirique ; devant deux formulations proposées pour une même question, plutôt que de se perdre en discussions sur les mérites et les dangers de chacune, il est préférable de les tester, c'est-à-dire de les poser à un petit échantillon et d'analyser les réponses que chaque forme suscite. Nous consacrerons, à la fin de ce chapitre, une section au pré-test du questionnaire. Mais c'est pendant toute la phase de conception et de rédaction qu'il faut avoir cette possibilité présente à l'esprit (cf. 4.6.).

Ceci dit, il reste possible d'énumérer, sans prétendre à l'exhaustivité, quelques remarques de portée générale.

Tout d'abord, le vocabulaire doit être simple. C'est évident, mais souvent difficile à

réaliser. Celui qui rédige le questionnaire est en général relativement cultivé ; il court, de ce fait, le risque soit d'utiliser un vocabulaire inhabituel qui ne sera pas compris par tous, ou alors, conscient de cette difficulté, de recourir à un langage artificiellement simplifié ou « populaire ». En principe, un des objectifs des entretiens non directifs préalables est précisément de connaître le vocabulaire utilisé par les sujets, pour savoir quels termes employer soi-même. Il faut toutefois se souvenir que les personnes interrogées considèrent en général l'enquêteur comme quelqu'un de cultivé et trouveraient bizarre de le voir employer certaines expressions populaires ou familières, même si elles font partie de leur propre langage.

Un autre danger lié au langage du chercheur provient de la familiarité de celui-ci avec les thèmes abordés dans l'enquête. Lorsqu'il est arrivé au stade de la rédaction du questionnaire, il y a en général plusieurs semaines, et parfois plusieurs années, qu'il réfléchit au problème, s'informe, en discute avec des spécialistes, qu'il lit ce qui a déjà été publié à ce propos. Il a donc accumulé à la fois une information générale qui le plus souvent dépasse largement celle de ses futurs sujets, un certain nombre de concepts spécifiques, et surtout un vocabulaire technique qui, pour être utile au spécialiste, a toutes les chances d'être différent de celui du reste de la population. En particulier, il pourra avoir tendance à utiliser certains termes courants dans un sens beaucoup plus précis que leur sens habituel. Un pré-test du questionnaire, en principe, devrait diminuer ce risque.

Souvent la définition de ce qu'on attend de l'enquêté est relativement simple et on peut la lui communiquer facilement. Dans d'autres cas, le concept précis élaboré par les spécialistes est plus complexe et n'a aucun équivalent dans la langue courante. Par exemple, le ministère du Travail effectue régulièrement une enquête sur la « population active », celle-ci étant définie de façon tout à fait précise. Or il est évident qu'on ne peut pas savoir si une personne est active ou non en lui posant directement la question, qui n'aurait pour elle au mieux qu'un sens très flou. Donner une définition précise comme on peut le faire pour les spécialistes est difficile. Cette définition est trop complexe, prévoit trop de cas différents pour qu'on puisse être certain qu'un sujet à qui on vient de la lire donne la réponse correcte. C'est pourquoi il a été pris un parti différent : la question est décomposée en plusieurs sous-questions, chacune correspondant à un des aspects de la définition, et suffisamment simple pour que chacun puisse répondre correctement sans difficulté, sans jamais utiliser le mot « actif ». C'est ensuite au codage qu'on fera les regroupements voulus.

Une question évidemment doit être comprise, et comprise de la même manière par tous. On s'efforcera donc d'éviter des termes vagues comme « souvent » ou « beaucoup » auxquels chacun peut donner une signification différente, en fonction de son cadre de référence. Il vaut mieux demander des estimations numériques, même imprécises, ou proposer des échelles de fréquence. Ce n'est toutefois pas toujours possible, certaines impressions étant très étrangères aux évaluations numériques, et imposer une telle évaluation risque d'apparaître comme complètement artificiel.

Parfois, l'emploi de ces termes ambigus peut être utile, non pas pour obtenir réellement une fréquence plus ou moins objective, mais pour savoir comment cette fréquence est ressentie. Par exemple, si on veut étudier les relations de voisinage et qu'on pose

les questions « est-ce qu'il vous arrive de vous rendre des services entre voisins ? », la réponse « souvent » ou « rarement » est probablement plus significative que « deux fois par semaine » qui n'aurait guère de sens, surtout si ce qui nous intéresse c'est l'insertion sociale, telle qu'elle est ressentie, et non la matérialité des faits.

Le sens de certains termes, même très simples, peut varier également selon la catégorie sociale ; le mot « livre » par exemple aura pour des personnes cultivées un sens beaucoup plus restreint que pour d'autres qui regrouperont sous ce terme tout ce qu'on peut lire et qui n'est pas le journal : autres périodiques, romans-photos, bandes dessinées, parfois feuilletons, etc. Dans la mesure où le terme ne peut en général pas être évité, on tournera la difficulté en le définissant clairement, éventuellement par énumération, ce qui présentera en outre l'avantage de rappeler certaines catégories, et donc d'éviter des erreurs dues à de simples oublis.

De même, le verbe « sortir », qu'on est facilement tenté d'utiliser dans les enquêtes sur les loisirs, peut avoir pour certains des connotations érotiques ou sentimentales qu'il n'aura pas du tout pour d'autres. Il faudra donc soit l'éviter, soit en préciser exactement la signification.

Un autre biais, très différent, consiste à formuler une question de manière à considérer certains présupposés comme acquis. Une plaisanterie classique consiste à demander à quelqu'un « s'il a cessé de battre sa femme », ce qui le met inévitablement dans l'embarras : qu'il réponde oui ou non, il aura reconnu qu'il battait sa femme. C'est un piège dans lequel on tombe facilement lorsqu'on rédige des questions, et souvent d'autant plus qu'on est mieux familiarisé avec le sujet : on considère certains points comme acquis et on pose une question qui n'a de sens que par rapport à eux.

Pour prendre un exemple très simple, la question « quel journal avez-vous lu hier ? » suppose qu'on a effectivement lu un journal la veille. Comme ne pas lire est souvent considéré comme dévalorisant, on peut s'attendre à ce que peu de personnes prennent l'initiative de répondre « aucun ». Une telle question est biaisée, et amènera à surestimer le nombre de lecteurs. Il faudrait commencer par demander si on a lu un journal (en précisant, évidemment, la signification qu'on veut donner à ce terme), puis ne demander lequel qu'à ceux qui auront répondu affirmativement.

L'utilisation de présupposés implicites peut être une méthode de manipulation, volontaire ou involontaire. Un des reproches adressés aux sondages d'opinion est précisément de poser des questions qui n'ont de sens que dans le cadre d'une certaine idéologie, et d'une conception de la société et de la politique et, de cette manière, de renforcer cette idéologie. Par exemple en posant des questions sur l'avenir d'un certain nombre d'hommes politiques, on admet que la politique est le fait d'individus et pas de groupes ou de classes, et que ce sont les mêmes individus qui seront encore importants dans quelques années. Nous reviendrons sur ce point à propos des questions d'opinion.

Ce même procédé peut être utilisé très consciemment pour faire admettre des points délicats qu'on avoue difficilement. Par exemple, les psychiatres, au cours des anamnèses, ne posent pas la question « vous êtes-vous masturbé ? » qui ne suscite que très peu de réponses affirmatives, mais ils demandent directement « à quel âge avez-vous commencé à vous masturber ? », question qui suppose une réponse positive à la précédente, et la considère comme normale.

Mais ce qui peut être légitime pour un psychiatre l'est beaucoup moins dans le cadre d'une enquête, qui ne correspond à aucune demande de la part du sujet interrogé. Il y a là une limite déontologique ; de plus il ne faut pas que le chercheur lui-même tombe dans des pièges de ce type, qui risquent de fausser sérieusement ses conclusions.

Les diverses précautions techniques, évidemment nécessaires, qu'on peut prendre pour éviter certaines erreurs, ne résolvent que très partiellement un des problèmes les plus difficiles de la méthodologie des enquêtes, celui de la formulation de questions ayant le même sens pour tous. Y réussir implique plusieurs conditions. Tout d'abord, nous l'avons vu, il faut non seulement que chaque terme soit compris, mais compris par chacun de la même manière. Ensuite, il faut que la question elle-même le soit, alors qu'elle peut être perçue de façon très variable, même si chaque mot en lui-même ne pose pas de problème. C'est toute la question des différences entre groupes sociaux, et de la manière dont on se représente ces différences.

Lorsque la variété de significations d'un même terme entre catégories sociales, ou entre régions, peut être considérée comme un simple problème de définition, comme dans les exemples que nous avons donnés ci-dessus, un changement de vocabulaire, ou l'explicitation de ce qu'on veut dire, peut éventuellement suffire. Mais la recherche d'un vocabulaire commun, même si elle aboutit, peut masquer des différences. Pour qui s'intéresse à la pratique de la lecture, savoir que pour certains un roman-photo est un livre, et pas pour d'autres constitue une information importante.

Dans certains cas, la relation entre le vocabulaire utilisé et ce que l'on veut étudier peut être encore plus étroite. Par exemple, certains ne feront entrer dans la « politique » que ce qui est directement lié à l'activité des partis ou du gouvernement, alors que pour d'autres cela comprendra tout ce qui touche à la vie sociale, considérant que l'implantation d'un centre commercial, les filières scolaires, ou l'attitude à l'égard de l'homosexualité sont tous des problèmes politiques. Dans cet exemple ce n'est probablement pas directement la catégorie sociale, mais plutôt la position politique elle-même qui déterminera le champ couvert par le mot « politique ». Si l'on pose une question où ce mot intervient, que ce soit comme adjectif ou comme substantif, la manière même dont la question sera comprise pourra faire partie de ce qu'on cherche à étudier.

Il est impossible de compter sur le refus des sujets de répondre à des questions trop éloignées de leur propre expérience pour éviter des erreurs ou des ambiguïtés de ce type. En effet, ces refus sont relativement rares : il est exceptionnel qu'un spécialiste pose des questions qui ne sont vraiment compréhensibles que pour une fraction de la population. Si une question est franchement refusée, le mal n'est peut-être pas trop grand : on sait où l'on en est. En revanche, ce qui est beaucoup plus fréquent, c'est que chacun projette sur la question la signification qui lui semble adéquate, sans qu'on ait les moyens de vérifier la concordance de ces significations.

Le souci, normal, d'appliquer un questionnaire identique à tous les membres d'une population se heurte donc à des obstacles sérieux, et surtout à des obstacles qu'il peut être dangereux de contourner trop vite. Même si l'on se garde de présupposés globalistes, en renonçant à dire quoi que ce soit sur la population dans son ensemble, on souhaite, en général, comparer les différentes catégories entre elles, et pour cela il faut

poser à tous les mêmes questions. Mais, ce faisant, on court le risque de réduire chaque groupe à ce qu'il a de commun avec les autres, ce qui a toutes les chances d'être très appauvrissant. Inversement, la juxtaposition de questionnaires, ou même de types d'enquêtes, différents, chacun bien adapté à sa catégorie, donnera une mosaïque dont il sera peut-être impossible de faire la synthèse.

4.5.2. Les questions de fait

4.5.2.1. Questions de fait et questions de connaissance

Nous avons vu que la distinction entre « questions de fait » et « questions d'opinion », malgré sa commodité, n'est pas toujours claire. En particulier, lorsqu'il s'agit de faits, il faut bien savoir si on s'intéresse aux faits « objectifs », c'est-à-dire tels qu'on pourrait en principe les connaître autrement que par ce qu'on nous en dit, ou si c'est la représentation ou la connaissance qu'en ont les sujets interrogés qui constituent l'information pertinente. Par exemple, on peut vouloir connaître le temps nécessaire à une personne pour aller à son travail, et juger que le lui demander directement est plus simple que de faire des mesures ; mais on peut aussi estimer qu'on agit en général non en fonction d'une mesure objective, mais de l'idée qu'on s'en fait, et donc que c'est elle qu'il faut connaître ; seule une question pourra nous apporter cette estimation subjective. Bien que, pour l'essentiel, les questions à poser soient les mêmes, certains chercheurs estimant même que la distinction n'a pas de sens, et que ce n'est jamais que des évaluations subjectives qu'on peut saisir par enquête, les précautions à prendre dans chaque cas, les recoupements qu'on pourra effectuer, ne seront pas les mêmes. Si ce qui nous intéresse est une information objective, on fournira au sujet toutes les aides possibles pour que sa réponse soit correcte ; dans le cas de la durée du déplacement jusqu'à son travail, on lui rappellera par exemple tous les trajets annexes, les temps d'attente, etc. Dans d'autres cas, on pourra l'inciter à se référer à des documents ou à d'autres traces, etc. Si au contraire c'est la représentation qu'on veut saisir, on se gardera évidemment de tout cela, mais on recourra peut-être à certains recoupements pour s'assurer que l'objet de la question a été bien compris.

Certains des faits sur lesquels peut porter une enquête sont par nature *uniques* : histoire personnelle ou professionnelle, appartenance passée ou présente à des associations, vote lors d'une élection particulière, accidents ou maladies, etc. Dans d'autres cas, peut-être plus fréquents, les faits qui intéressent le chercheur sont *répétitifs*, plus ou moins habituels ou réguliers : dépenses ou consommations diverses, utilisation du temps, déplacements, activités de loisirs, fréquentation d'équipements, lecture de journaux ou de livres, relations amicales ou contacts professionnels, etc. Un problème nouveau, propre à cette catégorie de comportements, vient s'ajouter aux autres, celui du choix de la période d'*observation*, problème plus complexe, et moins uniquement lié aux seuls problèmes de précision des mesures qu'il ne le semble au premier abord. Nous examinerons donc séparément les problèmes propres à l'étude de ces deux catégories de comportements.

4.5.2.2. Les événements uniques [1]

A première vue, une fois vérifiée la bonne compréhension du vocabulaire utilisé, le principal problème qui reste est celui d'éviter les *erreurs de mémoire*, qui peuvent se traduire, selon le cas, soit par de simples oublis, soit par d'autres déformations. Mais depuis quelques années, on s'est rendu compte que d'autres facteurs de distorsion interviennent ; facteurs qu'on peut rattacher à la relation enquêteur-enquêté, et au désir de ce dernier soit d'apparaître comme normal, acceptable, soit de se valoriser. Nous reviendrons plus en détail sur ce point dans le chapitre consacré aux aspects psychosociologiques de la situation d'enquête. Pour le moment, remarquons que ces distorsions peuvent prendre d'autres formes que celles de simples oublis, et que les erreurs peuvent être sélectives. On ne peut donc pas admettre non plus que les erreurs se feront nécessairement dans le sens d'une *sous-estimation* de la fréquence des comportements en question. Par exemple, lors d'une étude sur les voyages d'affaire, qui sont socialement très valorisés, on a constaté une surestimation de leur fréquence, certains sujets mentionnant des voyages que, d'après tous les recoupements, ils n'avaient pas faits. Dans le cas d'une enquête sur les hospitalisations, les séjours à l'hôpital pour des maladies « honteuses » ou socialement embarrassantes étaient plus sous-estimées que les autres.

Avoir voté ou non lors d'une élection, ou répondre sur les résultats scolaires des enfants, ont montré également des distorsions, dans le sens d'une plus grande acceptabilité sociale, qu'il est impossible d'attribuer à de simples oublis, bien que ceux-ci existent également.

Eliminer, ou plutôt diminuer, les distorsions plus ou moins conscientes, volontaires, attribuables à la relation enquêteur-enquêté, implique qu'on agisse sur celle-ci par le choix et la formation des enquêteurs, par le libellé des questions aussi (cf. 3.6. et 5.3.).

En ce qui concerne les erreurs de mémoire on a tenté, lorsque c'était possible, de proposer aux sujets des « aides » standardisées sous la forme de suggestion de chronologie (commencer par l'événement le plus récent, puis remonter dans le temps, ou inversement) ou de classifications. Leur effet semble avoir été minime.

En revanche, reprendre les questions après un certain intervalle (allant de 30 minutes à quelques jours) semble permettre de « récupérer » quelques événements oubliés. De même, faire reprendre les questions par un enquêteur expérimenté à qui on laisse une certaine latitude dans la manière de mener l'interrogation, s'adaptant à chaque cas, selon une technique proche de certaines interrogations critiques, ou de la méthode de Piaget, apparaît aussi comme très efficace.

4.5.2.3. Les événements répétitifs

Nous regroupons sous ce terme tous les événements ou les activités qui à la fois se présentent à plusieurs reprises au cours d'une certaine période, et qui sont relativement

---

1. Nous ne nous limitons pas ici aux événements uniques a sens strict, c'est-à-dire n'arrivant qu'une fois dans la vie d'une personne, mais nous regroupons sous ce terme tous les événements ou activités rares ou non habituels. Le fait que certains ne puissent pas être classés sans ambiguïté comme uniques ou répétitifs n'est pas important.

réguliers ou habituels. Il est donc nécessaire de décider de la durée de la période sur laquelle portera le questionnaire.

Interroger les sujets sur leurs comportements pendant une période de courte durée, allant en général de un jour à une semaine selon la fréquence supposée, présente l'avantage de réduire au minimum les erreurs dues à la mémoire. Si l'enquêteur demande de reconstituer l'ensemble des déplacements de la veille, par exemple, ou la fréquentation du cinéma durant la semaine écoulée, les déformations dues aux erreurs de mémoire ont des chances d'être très faibles.

Toutefois, les différents jours de la semaine n'étant en général pas équivalents du point de vue de nombreuses activités, et beaucoup d'entre elles étant réglées par un rythme hebdomadaire et non quotidien, on est obligé de prêter une attention particulière à l'*échantillonnage temporel,* et de faire en sorte que les échantillons interrogés chaque jour soient rigoureusement équivalents, à la fois en taille et en composition [1], ce qui rend très difficile, par exemple, l'échantillonnage par quota, où la latitude laissée aux enquêteurs est, en général, trop grande pour garantir que cette équivalence sera effectivement réalisée. Il faudrait aussi, idéalement, tenir compte des variations saisonnières et pour cela étaler l'enquête sur toute une année, ce qui, on le voit immédiatement, soulève des problèmes pratiques importants. En fait, seuls des organismes comme l'INSEE, qui dispose de moyens importants, et dont une des tâches est de rassembler des informations statistiques d'une grande précision, et d'étudier leurs variations dans le temps, peuvent se le permettre. Plusieurs des enquêtes de l'INSEE sont effectuées par vagues successives, tous les trimestres. En revanche, un organisme comme le SETRA (Service d'Etudes Techniques des Routes et Autoroutes, au ministère de l'Equipement) se contente, par ses recensements des déplacements urbains, d'une vague unique, choisissant de préférence, pour ces enquêtes, une période où ces déplacements sont importants, puisque ce sont, en partie au moins, sur ces périodes de pointe que sont prises les décisions d'aménagement et que sont élaborés les plans de circulation.

Lorsque la période d'observation est courte, on peut espérer que les informations recueillies seront à peu près exactes, ce qui amène certains de ceux qui conçoivent ou utilisent ces enquêtes, des économistes ou des ingénieurs, le plus souvent, à considérer que ce n'est que dans ces conditions qu'on dispose de « données objectives », et donc qu'il est nécessaire de s'y tenir, c'est-à-dire de poser des questions précises sur les comportements au cours d'une période aussi brève que possible. Toutefois, outre que, même dans ces conditions, l'exactitude rigoureuse des informations recueillies peut souvent être mise en doute, il faut aussi tenir compte du fait qu'elles restent très pauvres au niveau individuel. Les activités d'une journée déterminée caractérisent en général de façon très insuffisante les comportements d'une personne : tous les jours ne se ressemblent pas, les habitudes sont pour la plupart sujettes à fluctuations, certaines périodes sont exceptionnelles ; pour toutes les activités qui, pour une personne déterminée, ont

1. Cette dernière précaution est indispensable lorsque la périodicité hebdomadaire n'est pas la même dans différents groupes sociaux, comme les sorties du vendredi soir, par exemple, ou le fait de travailler le samedi ou le lundi, qui n'ont pas la même fréquence dans les différentes catégories socio-professionnelles. La manière de passer le mardi soir ou la journée du mercredi dépend de la présence d'enfants, etc.

une périodicité moyenne supérieure à la durée de la période d'observation, le fait qu'elles aient été effectuées durant cette période n'est guère significatif. On est amené à classer dans la même catégorie, par exemple celle des personnes ayant présenté un certain comportement la veille, aussi bien des personnes pour qui ce comportement est très fréquent, que d'autres pour qui il est exceptionnel, ou d'autres encore pour qui il est régulier, mais de périodicité très faible et qui se trouvent l'avoir présenté ce jour-là. C'est donc une manière très grossière, très peu discriminative, de caractériser les individus.

Ce n'est toutefois pas nécessairement un défaut grave. Si ce qu'on recherche c'est une estimation globale de la fréquence de ce comportement dans une population, ou même dans différentes sous-populations, l'information recueillie sur une période courte sera suffisante : statistiquement, les différentes variations s'annulent si l'échantillonnage, en particulier l'échantillonnage temporel, a été correct ; les estimations dans chaque sous-population seront satisfaisantes. Il faut simplement faire en sorte que la période d'observation ne soit pas trop inférieure à la périodicité moyenne du comportement étudié, afin de disposer d'un nombre suffisant d'observations pour atteindre la précision voulue.

Pour l'économiste qui cherche à estimer des grandeurs globales, ou pour l'ingénieur qui veut connaître, par exemple, des flux de déplacements, ces données sont en général suffisantes, et elles présentent bien l'exactitude dont ils ont besoin. Savoir que telle personne, qui traverse en voiture un certain carrefour aujourd'hui, le fait tous les jours ou seulement exceptionnellement, ne lui est guère utile. En ce qui concerne les voitures présentes le lendemain, leur nombre total aura peu varié, même si ce ne sont pas les mêmes.

En revanche, pour le sociologue ou le psychosociologue, plus intéressés, en général, comme nous l'avons vu, à analyser des relations qu'à obtenir de bonnes estimations de grandeurs globales, les observations de courte durée n'apportent pas les informations qui leur sont nécessaires : distinguer simplement les personnes qui sont allées au cinéma la semaine écoulée de celles qui n'y sont pas allées, est grossier et peu significatif. Celui qui veut analyser la fréquentation du cinéma, sa répartition dans différents groupes, sa signification culturelle ou les goûts de la clientèle, a en général besoin de distinctions plus fines. Certes, on peut admettre que, *en moyenne*, le groupe de ceux qui sont allés au cinéma le fréquente plus que le groupe de ceux qui n'y sont pas allés, mais si on veut pousser l'analyse en mettant en relation cette variable avec d'autres caractéristiques individuelles, c'est tout à fait insuffisant. Il faudrait disposer d'informations plus détaillées, donc portant sur une période plus longue, même si c'est au prix de la précision.

Pour obtenir de telles informations sur une longue période, une des méthodes utilisées, très lourde, mais qui donne des résultats satisfaisants, est celle du *carnet de comptes* employée surtout pour les études de budgets-dépenses ou de budgets-temps. Elle consiste à laisser à l'enquêté un carnet qu'on lui demande de remplir quotidiennement, ou même plusieurs fois par jour, en y inscrivant selon le cas toutes ses dépenses, ou toutes ses activités, ou plus simplement tout ce qui appartient à la catégorie sur laquelle porte l'enquête.

Cette méthode permet d'obtenir des informations à la fois précises et détaillées, particulièrement fiables, sur une période aussi longue qu'on le souhaite, et presque sans avoir recours à la mémoire. Par exemple, la première recherche sur les budgets de famille en France, effectuée par le CREDOC [1], a porté sur toute une année. C'est évidemment très astreignant pour les sujets, et il est nécessaire que, périodiquement, les enquêteurs reviennent les voir, s'assurent que les carnets ont été correctement remplis, posent des questions complémentaires pour préciser certains détails ou lever des ambiguïtés, et encouragent les sujets à continuer.

Cette méthode, évidemment très lourde et très coûteuse, présente le très gros avantage de fournir des informations à la fois riches et précises, même lorsque, comme c'est en général le cas, la durée de l'observation est moins longue. (Les études ultérieures sur les budgets-dépenses ont porté sur un mois, avec un questionnaire complémentaire recensant les dépenses de périodicité plus faible.)

On peut admettre que cette méthode diminue considérablement l'importance des oublis, rend psychologiquement difficile le mensonge conscient, et évite certaines erreurs d'estimation, comme la tendance à indiquer de préférence des nombres ronds. En revanche, on peut craindre que le fait de permettre aux sujets de prendre conscience de leurs comportements ne les modifie. On affirme souvent qu'un des moyens de limiter ces dépenses est d'en tenir un compte précis. D'autre part, avec le temps, il peut y avoir soit apprentissage de l'utilisation régulière du carnet, soit au contraire désintérêt croissant et donc oubli. Dans l'ensemble, les travaux méthodologiques effectués sur ces points suggèrent que la précision s'améliore après quelques jours et que, pour toutes les dépenses courantes, une durée d'observation d'un mois est satisfaisante.

Devant la lourdeur et le coût de ces méthodes d'enregistrement, et lorsque le thème de l'enquête n'exige pas une précision aussi grande, on se rabat en général sur des questions directes portant soit sur les comportements effectifs pendant une durée déterminée relativement longue, soit sur une estimation des fréquences « habituelles » de ces comportements.

Dans le premier cas, on traite les événements réguliers comme une suite d'événements uniques, avec tous les problèmes que nous avons évoqués ci-dessus. Si, en revanche, on parle d'« habitudes », on s'éloigne de la stricte description factuelle pour demander au sujet de porter des jugements sur le caractère régulier ou exceptionnel de certains de ces comportements, jugements qui ne découlent pas seulement de leur fréquence ou de leur caractère périodique. Les habitudes apparaissent, en effet, comme une norme, plus ou moins contraignante, mais en général susceptible d'écarts.

Cette ambiguïté de la notion d'habitude amène souvent les sujets, lorsque la norme et le comportement ne coïncident pas parfaitement, à donner des réponses complexes. Il y a ainsi certains comportements qu'on peut qualifier d'« habituels rares » ou inversement d'« exceptionnels fréquents ». C'est le cas, par exemple, de la marche à pied : comme elle est socialement valorisée, beaucoup de personnes affirment effectuer « normalement » ou « habituellement » à pied certains déplacements ; ils suggèrent ainsi que presque chaque fois qu'ils font ces trajets, ils les font à pied. Pourtant, si on

1. Il est intéressant de noter que, bien que la période d'observation semble très longue, les auteurs la trouvent encore insuffisante : « ... on ne connaît qu'une courte tranche de la vie d'un ménage, non sa totalité. »

prolonge l'entretien, on s'aperçoit que ces mêmes personnes prennent leur voiture quand il pleut, quand elles sont chargées ou fatiguées, etc., ce qui fait que, finalement, c'est leur « comportement normal » qui, du point de vue de sa fréquence, est exceptionnel. Répondant à une seule question concernant leurs habitudes, ces mêmes personnes donnent une estimation très exagérée de la proportion de leurs déplacements effectués à pied.

On rencontre évidemment le phénomène inverse pour les comportements dévalorisés, comme la lecture de certains journaux du type *France-Dimanche* ou la consommation de chocolat dont « normalement » on s'abstient, sauf quand une occasion se présente ; le problème, pour celui qui veut décrire ces comportements, est de savoir si ces occasions « exceptionnelles » ne se trouvent pas être fréquentes.

Les réponses aux questions portant directement sur les habitudes ne peuvent donc pas être acceptées telles quelles, comme des descriptions de comportements effectifs. Mais, inversement, demander à des sujets de telles descriptions sur une période suffisamment longue pour que l'information soit riche met en jeu les problèmes de mémoire et aussi, parfois, le caractère exceptionnel de cette période. L'expérience montre que les questions sur les habitudes donnent des résultats qui conduisent à des analyses intéressantes. On ne peut donc pas en conclure, comme le font hâtivement certains critiques, que les sujets répondent n'importe quoi. En revanche, il est très difficile de savoir dans chaque cas quels sont les biais spécifiques qui peuvent s'introduire.

Lorsqu'on demande au sujet, quelle que soit la forme de la question, d'indiquer une fréquence, ou éventuellement une proportion, il faut prendre garde à l'attraction des nombres ronds ; en général cela n'est pas très important dans la mesure où il est rare qu'on tienne à connaître des grandeurs à une unité près. Le phénomène peut diminuer la précision des estimations, il ne semble pas qu'il introduise une distorsion systématique ; tout au plus viendra-t-il renforcer une tendance à la surestimation ou à la sous-estimation.

Le choix de la durée de la période d'observation a donc des effets sur la nature des informations disponibles au niveau individuel. Il a aussi des conséquences, moins apparentes, sur les relations qu'on peut mettre en évidence entre des comportements. Supposons que, au cours d'une recherche sur les loisirs, nous nous intéressions à la fois à la fréquentation du cinéma et à celle du théâtre. Si les questions posées portent sur une période courte, la journée ou peut-être la semaine, on observera une relation *négative* entre les deux : ceux qui sont allés au cinéma seront rarement allés au théâtre, et inversement. Si l'on allonge la période, l'étendant à quelques mois, on voit la relation s'inverser, devenir *positive* : ce sont les mêmes personnes qui fréquentent à la fois ces deux types de spectacles, ce qu'on peut interpréter comme des manifestations différentes d'un même goût, ou par des habitudes culturelles qui amènent à fréquenter les deux. En revanche, entre la fréquentation du théâtre et celle de spectacles sportifs, la relation reste négative, bien que faible. Dans le premier cas, ce sont les mêmes personnes qui présentent les deux comportements, mais les contraintes de temps les empêchent de pratiquer les deux au cours d'une même courte période. Dans l'autre cas, ce sont deux groupes différents qui fréquentent les deux types de spectacles. Enfin, si nous prenons comme troisième exemple, toujours dans le domaine des loisirs, la promenade en voi-

ture et le pique-nique, on voit que la relation reste toujours positive, même sur une période courte : il n'y a pas substitution, mais complémentarité [1].

On peut trouver sans difficulté des exemples dans d'autres domaines : les consommations de viande de bœuf ou de veau présentent, comme le cinéma ou le théâtre, une relation négative sur une courte période, et positive lorsque la période s'allonge, exprimant les différences de consommation de viande en général dans la population. En revanche, la relation entre le café et le sucre reste positive, et asymétrique puisqu'on peut prendre du sucre avec autre chose que du café.

4.5.3. Les questions d'opinion

4.5.3.1. Problèmes généraux

Passer des questions de fait à ce que nous avons appelé des questions d'opinion, nous fait entrer dans un domaine beaucoup plus flou. Déjà, nous l'avons vu, les questions portant sur des faits sont loin d'être aussi simples, univoques et dépourvues d'ambiguïté qu'on pourrait le croire au premier abord. Nous avons vu au chapitre précédent un certain nombre des problèmes que pose ce type de questions ; nous les retrouverons, se posant à peu près de la même manière, à propos des questions d'opinion, avec, en plus, les difficultés spécifiques qu'entraîne l'absence de référent objectif.

En principe, la validité des questions de fait peut être évaluée. Dans certains cas au moins on peut comparer les réponses obtenues avec des informations provenant d'autres sources, considérées comme plus objectives, et vérifier ainsi si les questions suscitent bien les réponses souhaitées. Pour les questions d'opinion une telle validation externe est impossible par définition. On doit se contenter de validations internes, comme l'appréciation de la cohérence des différentes réponses obtenues sur un même thème, ce qui est évidemment insuffisant comme validation : cette cohérence peut avoir été suscitée par la situation d'enquête, par le désir du sujet interrogé de donner de lui-même l'image d'une personne rationnelle, dans une situation où on l'amène à rapprocher des points auxquels il aurait peut-être, jusqu'alors, réfléchi séparément. La cohérence avec les comportements n'est pas suffisante non plus : de nombreuses recherches ont montré que l'accord entre attitudes mesurées et comportements est rare, soit que l'indicateur d'attitude utilisé ne soit pas bon, soit que les comportements aient d'autres déterminants que l'attitude étudiée : d'autres attitudes, des contraintes diverses, les informations disponibles, peuvent aussi bien affecter les comportements. Quelles que soient les précautions méthodologiques qu'on puisse prendre, il reste donc toujours une incertitude quant à la validité des questions d'opinion.

Lorsqu'on pose une question sur un comportement, sa formulation peut être complexe, mais, en principe, on devrait toujours être capable de communiquer au sujet de

1. Dans ce cas particulier, la relation n'est pas symétrique : la sortie en voiture est en général une condition nécessaire, mais non suffisante, du pique-nique. Il ne faudrait donc pas analyser ces données au moyen d'un coefficient de corrélation ou d'association, qui suppose une relation symétrique, mais vérifier l'existence d'une implication. On peut d'ailleurs observer la même dissymétrie entre la fréquentation du cinéma et celle du théâtre, bien que l'interprétation de l'implication ne soit pas la même (sur l'analyse en relations asymétriques voir 7.2.1.2., a et 7.3.2.).

quoi on veut qu'il parle. Lorsque la question porte sur une opinion, en revanche, rien ne nous assure que cette opinion existe effectivement, même pas le fait que le sujet a répondu à la question. Non seulement certaines des personnes interrogées peuvent n'avoir jamais réfléchi au problème dont il est question, donc ne pas avoir vraiment d'opinion à son propos, mais il se peut aussi que leur manière de concevoir le thème soit toute différente de celle qui sous-tend la question et qui a orienté sa formulation ; leur cadre de référence, les informations sur lesquelles ils s'appuyent, peuvent aussi différer considérablement de l'un à l'autre, modifiant la signification qu'ils attribueront à la question, et donc celle de leur réponse. Dans certains cas, on peut dire que c'est la question qui crée l'opinion : sans elle cette opinion serait restée, chez certaines personnes au moins, quelque chose de vague, de confus, d'inorganisé ; la question lui fournit une structure.

Des réflexions de ce type amènent certains à critiquer radicalement les enquêtes qui portent sur autre chose que sur des faits ; on objecte qu'on demande à des personnes de se prononcer sur des points auxquels elles n'ont peut-être jamais réfléchi, et à propos desquels elles ne sont que très insuffisamment informées. Quelle valeur peut avoir leur réponse dans ces conditions ? De telles critiques ne sont pas exemptes de considérations politiques ou morales : ce qui est en jeu, c'est le poids politique qu'il faut accorder à l'opinion des non-spécialistes ; nous n'en parlerons pas ici. Le problème technique qui se pose n'est pas de savoir s'il faut accorder une valeur quelconque à l'opinion d'une personne jugée incompétente, mais de déterminer ce que signifient ses réponses. On peut penser que celles-ci sont instables, très sensibles à la formulation des questions et au contexte du questionnaire ; mais cela ne fait probablement que reproduire l'instabilité de l'opinion elle-même qui, selon les circonstances et le contexte, sera façonnée, modifiée, et pourra se traduire par des actions très diverses.

Un hebdomadaire a récemment publié une enquête au cours de laquelle on demandait aux personnes interrogées de refaire le budget de l'Etat, en répartissant selon leurs préférences les sommes attribuées à chaque ministère. C'est un bon exemple de question technique posée à des personnes en principe incompétentes : peu d'entre elles, vraisemblablement, savaient exactement ce que recouvrent les différents postes du budget, quelles sont les attributions précises de chaque ministère, ou quelles sont les contraintes qui pèsent sur cette répartition. De plus, comme il s'agit d'un problème rarement discuté dans ces termes dans la presse, les sujets n'avaient, pour la plupart, guère eu l'occasion d'y penser auparavant, et donc d'apprécier toutes les conséquences de leur choix [1]. En tant que guide pour l'élaboration de la Loi de Finances, cette enquête sera probablement jugée sans utilité immédiate. Néanmoins, en tant qu'expression de choix politiques plus fondamentaux, les réponses obtenues sont certainement significatives de grandes options, ce qui justifie l'enquête. Toutefois, ces significations risquent d'être variables selon ce que chacun met derrière les termes généraux utilisés : on confond ainsi des réponses qui sont apparemment semblables mais qui, en tant qu'indicateurs de positions plus générales, risquent d'être, en fait, profondément différentes.

1. Cette enquête comportait quelques questions cherchant à vérifier le niveau d'information des personnes interrogées. Il aurait été intéressant de savoir si les répartitions proposées par les personnes bien informées différaient de celles des autres, et comment. Mais cette analyse, si elle a été effectuée, n'a pas été publiée.

Ces divergences ou ces ambiguïtés, attribuables à des niveaux de réflexion préalable et d'information différents, sont atténuées lorsque les problèmes sont socialement posés, indépendamment de toute enquête. Lorsqu'un problème est discuté, lorsqu'on en parle dans les journaux et à la télévision, lorsqu'il fait l'objet de conversations, il se produit probablement un consensus non sur la meilleure solution à y apporter, mais sur une certaine manière de le poser, sur la gamme des options possibles, etc. Un questionnaire qui s'appuye sur ce consensus a des chances d'être bien compris. Le cas extrême est celui des référendums ou des élections : l'enquête n'a alors qu'à reproduire les questions posées officiellement. C'est pour cela, d'ailleurs, que les succès impressionnants de certains sondages pré-électoraux ne constituent pas une véritable validation des enquêtes *en général* : les conditions sont trop favorables, et ne se retrouvent pas dans les autres types d'enquêtes [1].

Certains objets d'enquête n'ont pas subi une telle élaboration sociale : on n'en parle guère, et ils ne correspondent pas immédiatement à des problèmes de la vie quotidienne, ou à des événements précis. C'est le cas, par exemple, lorsqu'on veut étudier des attitudes de portée générale, et non l'opinion à l'égard de points précis. De telles variables sont conçues par le chercheur lui-même, à la suite de considérations théoriques, par inférence à partir d'observations diverses, ou par abstraction à partir de notions du sens commun. La « mobilité », utilisée dans certaines enquêtes sur les déplacements, le « modernisme », l'« ethnocentrisme », constituent des exemples de telles attitudes, dont rien ne nous assure a priori qu'elles existent. L'affirmation de l'existence de ces variables constitue elle-même une hypothèse qui doit être vérifiée.

La distinction que nous avons faite entre questions de fait et questions d'opinion est commode, mais elle est difficile à maintenir dans tous les cas. Considérer qu'une proposition est l'affirmation d'un fait, ou n'est que l'énoncé d'une opinion, peut être déjà significatif. Prenons la phrase « les Etats-Unis défendent la liberté partout dans le monde » : pour certains ce sera l'énoncé, vrai, d'un fait ; pour d'autres, ce sera une opinion, fausse. Répondre qu'on fait quelque chose « souvent » ou « rarement » peut constituer aussi un jugement de valeur et correspondre en fait, chez des personnes différentes, aux mêmes fréquences. La distinction entre questions de fait et d'opinion porte donc plus sur l'intention du chercheur, sur la manière dont il utilisera ou interprétera les réponses obtenues, que sur les questions elles-mêmes.

4.5.3.2. La forme des questions et des réponses

Lorsqu'on utilise des questions fermées, ce qui est le cas le plus fréquent, la forme des questions, et celle des réponses proposées peut être très variable. Mentionnons les plus couramment utilisées.

a) On énonce une opinion, précédée éventuellement d'une phrase introductive comme « voici une opinion qu'on entend parfois », et on demande au sujet s'il est d'accord

2. Sur un point toutefois, on peut tirer de ces succès des conclusions plus générales : ils montrent que les procédés d'échantillonnage utilisés sont satisfaisants, qu'on est effectivement capable de constituer des échantillons représentatifs de la population des électeurs, par une combinaison de sondage aréolaire et de la méthode des quota.

ou non avec elle. Par exemple, dans un questionnaire sur la représentation du psychologue, on a proposé :

« Le psychologue est indispensable à la bonne marche de l'entreprise » :
— d'accord,
— pas d'accord.

Eventuellement, on pourra atténuer ce que ces réponses ont de trop catégorique en proposant des réponses intermédiaires :

— tout à fait d'accord,
— plutôt d'accord,
— plutôt pas d'accord,
— pas du tout d'accord.

Parfois, on simplifie en laissant tomber les deux réponses extrêmes, les réponses intermédiaires étant plus généralement acceptées.

Lorsque l'opinion peut apparaître comme l'énoncé d'un fait, l'alternative vrai/faux peut aussi être proposée. Mais cela présente l'inconvénient de suggérer au sujet qu'il y a une « bonne réponse », ce qui risque d'introduire chez lui une attitude exagérément critique et prudente.

b) On présente la même opinion sous forme interrogative, la réponse étant alors « oui » ou « non ». Ainsi, la question ci-dessus aurait pu être formulée :

« Le psychologue est-il indispensable à la bonne marche de l'entreprise ? »

ou

« A votre avis (ou « selon vous »), le psychologue est-il indispensable à la bonne marche de l'entreprise ? »

c) On propose plusieurs opinions sur le même thème, et on demande aux personnes interrogées laquelle correspond le mieux à leur propre position. Par exemple, toujours sur la représentation du psychologue :

« Avec laquelle de ces opinions êtes-vous le plus d'accord :
— Le psychologue est indispensable à la bonne marche de l'entreprise.
— Le psychologue rend des services dans une entreprise, mais on peut s'en passer.
— Le psychologue est totalement inutile dans une entreprise.
— Le psychologue ne sert qu'à masquer les vrais problèmes de l'entreprise. »

Selon le cas, on pourra demander de ne choisir qu'une seule réponse, d'en choisir plusieurs, de les ordonner, etc.

La première formule, l'énoncé d'une opinion à propos de laquelle on exprime son accord ou son désaccord, est la plus souple. Elle évite la forme interrogative, qui risque d'être source de malentendus et d'incompréhensions, surtout lorsque l'opinion est énoncée sous forme négative. Elle permet également d'introduire des nuances dans les réponses, sous la forme de degrés d'accord ou de désaccord, ce qui est plus difficile lorsque les réponses sont « oui » ou « non ».

Une limite de ces deux formes de questions, c'est que si on connaît relativement bien la position de celui qui a répondu « oui » ou « d'accord », on sait moins clairement ce que signifie la réponse négative. Dans l'exemple que nous avons pris, on ne sait pas si la personne, qui n'admet pas que le psychologue soit indispensable, estime qu'il est sim-

plement utile, ou qu'il est complètement inutile, ou encore qu'il est dangereux. Si on veut lever cette ambiguïté, il est nécessaire de poser d'autres questions. C'est là que la troisième forme, le choix entre plusieurs opinions, peut être utile. En imposant de n'en choisir qu'une seule on les rend mutuellement exclusives, ce qui évite certaines incohérences (il faut évidemment s'assurer, au cours du pré-test, qu'elles sont vraiment reconnues comme exclusives par les sujets, sinon on impose une contrainte extérieure qui fausse la signification des réponses). D'autre part, en présentant ainsi simultanément différentes positions possibles, chacune prend son sens par rapport aux autres, ce qui contribue à les rendre moins équivoques.

En ce qui concerne les réponses à proposer, plusieurs problèmes se posent. En premier lieu, décider si on proposera explicitement au sujet la réponse « je ne sais pas » ou « sans opinion », où si l'on se contentera de noter lorsqu'un sujet déclare ne pas avoir d'opinion. Dans le premier cas, le nombre de « sans opinion » a toutes les chances d'être plus élevé. Le choix dépend de la signification qu'on accorde à ce type de réponse. Si l'on admet que ne pas avoir d'opinion est une position aussi intéressante à connaître que les autres, on la proposera aux sujets sur le même plan que les opinions proprement dites. C'est le plus souvent le cas, et l'analyse des résultats recueillis montre que le nombre des « sans opinion » peut varier considérablement d'une question à l'autre dans un même questionnaire, et n'être pas non plus le même, pour une question donnée, dans différentes catégories sociales. C'est donc une indication importante, qu'il est nécessaire de recueillir comme les réponses proprement dites, et dont il faut tenir compte dans l'interprétation des résultats.

Néanmoins, il arrive que le chercheur estime qu'il est peu vraisemblable que quelqu'un n'ait aucune opinion sur un certain problème, et que l'absence de réponse exprime plutôt l'absence de réflexion, ou masque un refus de répondre pour des raisons diverses. On cherchera alors à diminuer le nombre des « sans opinion » en ne proposant pas cette possibilité dans la liste des réponses, tout en prévoyant, en revanche, une rubrique « refus de répondre », qui sera ainsi distinguée des « sans opinion ».

Il arrive également qu'une question ne nous intéresse pas pour elle-même, mais uniquement comme élément d'un indicateur complexe, par exemple une échelle d'attitude construite à partir des réponses à plusieurs questions (cf. 7.2.1.2.). Dans ce cas, le fait qu'un sujet n'ait pas répondu à l'une des questions rend inutilisables ses réponses à toutes les autres questions qui interviennent dans l'échelle, sur laquelle il devient impossible de le situer. Devant ce risque de perte d'information importante, il est peut-être préférable d'inciter fortement le sujet à faire l'effort de répondre.

Un des avantages de l'alternative « d'accord/pas d'accord » comparée à « oui/non », c'est qu'il est plus facile de la nuancer, de trouver des termes intermédiaires comme ceux que nous avons cités plus haut. Toutefois la précision supplémentaire apportée par ces échelles fixes est souvent illusoire : on s'aperçoit à l'analyse qu'il est préférable de regrouper des réponses voisines, et de les ramener à deux catégories seulement. Néanmoins, avoir proposé quatre réponses, et non deux, présente plusieurs avantages. Tout d'abord, les personnes interrogées peuvent avoir l'impression, avec deux réponses possibles seulement, qu'on les force à prendre des positions tranchées qui ne sont pas les leurs, et qu'on ne s'intéresse pas vraiment au détail de ce qu'elles

pensent. Ensuite, ayant proposé plus de deux réponses, on peut choisir le regroupement qu'on juge le meilleur, et pas nécessairement mettre ensemble les deux premières réponses d'une part, les deux autres d'autre part. Par exemple, si une proportion importante des sujets a choisi « tout à fait d'accord », on pourra regrouper les trois autres réponses, qui constitueront le groupe de ceux qui expriment des réserves. On évite ainsi de conserver, au cours de l'analyse, des catégories ne comprenant que très peu de sujets.

De même qu'il faut se demander si on va prévoir une réponse « sans opinion », il faut se demander si on va prévoir une catégorie intermédiaire ou neutre. Les avis, et les pratiques, des spécialistes sont très partagés sur ce point. Certains estiment que si l'on pense qu'une position neutre peut exister, il faut qu'elle puisse s'exprimer de la même manière que les autres. Mais on peut aussi craindre d'offrir ainsi une échappatoire à ceux qui ne feront pas l'effort de réfléchir. D'autre part, il est parfois difficile de distinguer la neutralité qu'on peut aussi considérer comme une position, au même titre que les autres, de l'absence d'opinion. Si on propose une telle réponse, on peut s'attendre que ce soit elle que choisiront ceux qui, autrement, se déclareraient sans opinion ou refuseraient de répondre.

La formulation verbale explicite d'une position neutre est, en général, difficile : proposer par exemple « à la fois d'accord et pas d'accord » ou « ni d'accord ni pas d'accord » n'aura évidemment pas le même sens, même si les deux expressions peuvent, sans incohérence apparente, être placées en position médiane. Le mot « neutre », explicitement proposé, est ambigu et souvent considéré comme un synonyme de l'absence d'opinion.

Une manière d'éviter ces problèmes consiste à proposer une échelle de réponses numérique, dont seule la signification des extrêmes est explicite :

— 2	— 1	0	1	2
ou 1	2	3	4	5
pas du tout d'accord			tout à fait d'accord	

La première échelle, en opposant des valeurs négatives et positives, et en marquant la position médiane par le zéro, met plus que la seconde l'accent sur une certaine organisation de l'opinion, ce qui peut aider les sujets, mais aussi, peut-être, les orienter. On pourra aussi proposer une échelle de la forme :

tout à fait d'accord — pas du tout d'accord

Le sujet met simplement une croix dans la case correspondante, ce qui évite d'éventuelles connotations attribuées aux chiffres.

Cette manière de faire évite la plupart des difficultés de la formulation verbale des

positions intermédiaires. On peut augmenter autant qu'on le souhaite le nombre d'échelons, et selon qu'on en proposera un nombre impair ou pair, on aura introduit ou non un point neutre dans l'échelle. Toutefois, pas plus qu'une échelle verbale, une échelle numérique ne garantit contre les risques de distorsion dans les évaluations par les différents sujets.

Les trois formes de questions que nous avons distinguées sont les plus courantes, mais il y en a d'autres, adaptées à des problèmes spécifiques. Nous en verrons quelques-unes plus loin.

4.5.3.3. Le libellé des questions

Plus encore que dans les autres points abordés dans ce livre, nous nous trouvons dans le domaine du tour de main et de la méthodologie négative. Nous pouvons énumérer des précautions à prendre, signaler des erreurs à éviter : cela ne suffit pas pour indiquer comment formuler de bonnes questions.

Plus que dans le cas des questions de fait, il est nécessaire de s'assurer que l'objet de la question a bien une réalité et une unité suffisantes pour tous les membres de la population étudiée. Une question portant globalement sur « la politique actuelle du gouvernement », par exemple, pourra avoir un sens très clair pour certains, alors que pour d'autres il faut en distinguer différents aspects, ce qui les rend incapables d'un jugement d'ensemble. Cela est d'autant plus gênant que cette difficulté risque d'être en relation avec l'opinion qu'on cherche à saisir : il est possible que ceux qui sont soit très favorables, soit très hostiles au gouvernement, puissent répondre sans difficulté à une proposition aussi générale, alors que certains de ceux qui ont une position intermédiaire voudront répondre qu'ils approuvent certains aspects de cette politique et pas d'autres. Même l'introduction, parmi les réponses proposées, de termes nuancés ou neutres, ne permettra pas l'expression de ces positions. Il faut admettre que « la politique du gouvernement » n'est pas un objet d'opinion pour tous, et découper la question, demandant aux personnes interrogées de se prononcer séparément sur la politique extérieure, la politique économique, la politique sociale, etc. Cette décomposition devrait se faire non a priori mais à la suite d'une analyse de jugements portés de façon plus spontanée, dans des entretiens non directifs, par exemple, ou par analyse du contenu de la presse.

a) *Il faut veiller à la structure logique de la question.* Les *négations*, en particulier dans les phrases interrogatives, et surtout les doubles négations, peuvent être sources d'ambiguïté, d'incompréhension et d'erreur. Il ne suffit pas de s'attacher à la structure de la question elle-même, il est nécessaire d'y inclure aussi les différentes réponses possibles. Il faut pouvoir répondre indifféremment « oui » ou « non » à une question sans qu'une des réponses pose plus de problèmes que l'autre.

Une proposition telle que « Il ne faut pas que le psychologue ait beaucoup de pouvoir dans l'entreprise » est acceptable si on demande aux sujets d'exprimer leur accord ou leur désaccord. Mais répondre devient très difficile si on la met sous forme interrogative : « Ne faut-il pas que le psychologue n'ait pas beaucoup de pouvoir dans l'entreprise ? » devient incompréhensible, et il est impossible d'y répondre par « oui » ou par

« non ». Même la forme plus simple « Faut-il que le psychologue n'ait pas beaucoup de pouvoir dans l'entreprise ? » risque d'être mal comprise, et la réponse pourra être ambiguë.

Dans un questionnaire, il faut qu'on puisse répondre « non » à toute question, que toute opinion puisse être niée. La négation est donc potentiellement toujours présente, même si la phrase elle-même est entièrement positive. Les problèmes liés à la négation sont donc importants, et apparaissent sous des formes très diverses. Ce sont eux, en fait, qui sont sous-jacents aux difficultés que nous allons examiner maintenant.

b) *Il faut éviter qu'une même réponse puisse être donnée pour des raisons trop différentes.* Par exemple, une opinion politique intrinsèquement modérée pourra être refusée par tous ceux qui ont des positions plus radicales, qu'ils soient de droite ou de gauche. On aura donc ainsi confondu des personnes ayant des positions très différentes, voire opposées. Dans ce cas, comme cela arrive souvent, une des réponses (ici l'approbation de l'énoncé) est acceptable, mais pas l'autre, à cause de son ambiguïté. L'examen critique d'une question doit porter non seulement sur son contenu, mais aussi sur la signification de chacune des réponses.

Par exemple, au cours d'une enquête portant sur l'attitude à l'égard de l'intervention de l'Etat dans des domaines pouvant être considérés comme appartenant à la vie privée, on avait proposé l'opinion suivante :

« Il est normal que certaines vaccinations soient obligatoires, mais il faut que des dispenses puissent être facilement accordées. »

Cette question avait été conçue pour exprimer une position moyenne, et effectivement elle était approuvée par tous ceux qui avaient une attitude modérée sur ce point ; leur réponse n'était pas ambiguë. En revanche, on s'en est aperçu à la fois par les commentaires des sujets au cours de la passation et lors de l'analyse des relations entre cette question et d'autres, cette même proposition a été repoussée à la fois par les partisans convaincus de la vaccination obligatoire, qui estimaient que les dispenses doivent rester tout à fait exceptionnelles, et par les adversaires, pour qui même la possibilité de dispense ne rendait pas l'obligation plus acceptable [1].

Dans cet exemple, la question était intrinsèquement modérée, et a toujours été comprise comme telle. D'autres peuvent être aussi interprétées comme simplement une position minimale, qui n'exclut pas des opinions plus extrêmes. Par exemple, la question :

« Souhaitez-vous la nationalisation des très grosses entreprises ? »

peut être comprise de deux manières :

« Souhaitez-vous la nationalisation des très grosses entreprises et d'elles seulement ? »

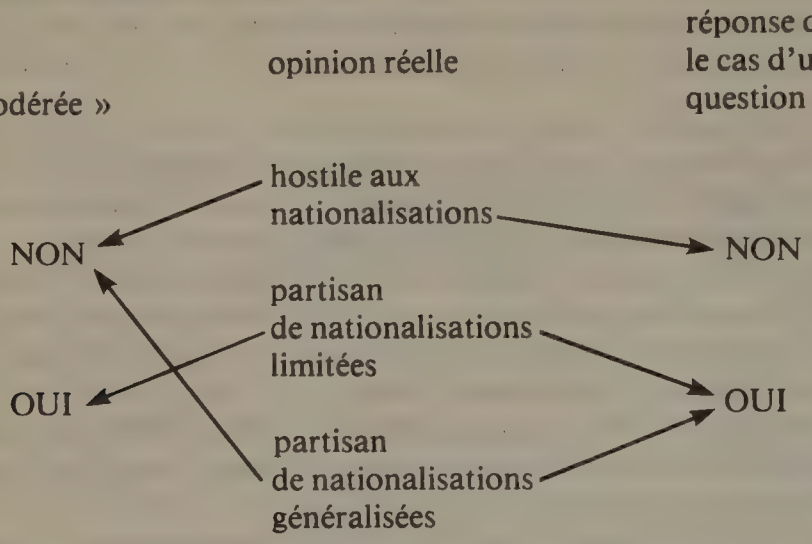
« Souhaitez-vous la nationalisation des très grosses entreprises, sans vous prononcer sur d'autres ? »

La première interprétation est celle que nous avons appelée « intrinsèquement modérée » ; comme dans l'exemple précédent, pourraient répondre « non » aussi bien ceux qui sont hostiles à toute nationalisation, que les partisans de nationalisations plus

1. Cette question peut aussi être critiquée parce qu'elle contient deux idées, ce qui, nous le verrons plus loin, est à éviter.

poussées. La seconde interprétation, en revanche, « minimale », suscitera des réponses moins ambiguës, puisque ces derniers pourront répondre « oui », comme ceux qui ne souhaitent que des nationalisations limitées.

La correspondance entre les réponses recueillies et l'attitude réelle peut être représentée par le schéma suivant :



On voit que seule une question « minimale » permet une inférence correcte sur l'opinion « réelle » ; certes, elle confond aussi deux groupes, mais ce sont des groupes d'opinions voisines, et non opposées.

L'ambiguïté entre les deux interprétations pourra être levée soit par le libellé même de la question, qui en préciserait la signification, soit par le contexte du questionnaire. Si on propose une série de questions de la même forme, portant sur l'opportunité de différentes nationalisations envisageables, les personnes interrogées comprendront plus facilement que c'est dans le sens minimal que la question leur est posée.

En termes généraux, on peut dire que ces problèmes sont liés à l'ambiguïté de la négation dans le langage et dans le raisonnement naturels. Une proposition peut être très claire et comprise sans difficulté : cela n'entraîne pas que sa négation le soit. Dans certains cas, nier une proposition, c'est en affirmer le contraire ; dans d'autres, c'est accepter tout ce qui n'est pas elle. Nous retrouverons ces problèmes plus loin (cf. 7.2.1.2.4.), lorsque nous examinerons la structure logique sous-jacente à l'analyse hiérarchique. Ce que nous avons appelé les questions « minimales » sont celles qui permettent la construction d'échelles hiérarchiques.

c) *Il ne faut en aucun cas introduire deux idées dans la même question.* Par exemple, l'énoncé :

« Il est regrettable que tellement d'œuvres importantes n'aient pas pu paraître à cause de la censure »

contient en fait deux idées : a) tellement d'œuvres importantes n'ont pas pu paraître à cause de la censure, et b) c'est regrettable.

Que devra répondre celui qui juge que, en effet, la suppression d'œuvres importantes est regrettable, mais qui, en même temps, pense que cela ne s'est pas « tellement » souvent produit ? Logiquement une proposition, formée par la conjonction de deux autres, est fausse si une au moins des deux propositions l'est. Il suffirait donc qu'une personne soit en désaccord avec une des deux affirmations pour qu'elle réponde « non ». Ce n'est pas ainsi qu'on raisonne d'habitude, et ce n'est probablement pas ce que souhaitaient les auteurs de la question.

Comme il arrive souvent, répondre à cette question ne présente aucune difficulté pour certains sujets ; ici, ce sont ceux qui sont les plus hostiles à la censure qui répondront « d'accord » sans problème. Mais pour tous les autres, répondre sera à peu près impossible. Ou alors, constatant que c'est une proposition hostile à la censure, une personne favorable répondra « pas d'accord », sans entrer dans le détail de la signification de la question.

Ne pas s'apercevoir que plusieurs idées sont incluses dans un même énoncé est souvent, ici comme dans bien d'autres exemples d'erreurs considérées comme « techniques », la conséquence de présupposés du chercheur. Dans notre exemple, celui-ci était vraisemblablement un intellectuel libéral qui considérait comme acquis que la censure a eu des effets néfastes ; la seconde partie de la phrase a donc dû lui apparaître simplement comme l'énoncé d'un fait évident, sur lequel on demande aux sujets de porter un jugement moral, et non de vérité.

d) *Certains termes chargés d'affectivité, de jugements de valeur, de connotations diverses modifient le sens de la question, et donc les réponses.* Pour reprendre un exemple célèbre, les formulations « Faut-il autoriser... ? » ou « Faut-il interdire... ? » ne suscitent pas des réponses équivalentes. Mise à part la tendance à l'acquiescement, dont nous parlerons plus loin, on rencontrera des personnes qui repoussent l'idée d'interdire quelque chose, mais qui en même temps s'indignent que cette même chose puisse être autorisée. C'est que de telles questions ne mettent pas exclusivement en jeu l'attitude à l'égard de l'objet de l'éventuelle interdiction ou autorisation, mais aussi des attitudes plus générales comme l'autoritarisme, la relation aux autorités, la représentation de l'Etat et de soi-même comme plus ou moins libéral, etc.

Devant la constatation de ces difficultés on se demandera tout naturellement quelle est la « bonne question », celle qui suscitera la « vraie réponse ». En fait on peut admettre que toute formulation est biaisée, et que se poser le problème de la « bonne » forme n'a guère de sens. Certes, on peut imaginer une question peut-être plus neutre, comme « Faut-il qu'on puisse... ? ». Certains biais seraient peut-être éliminés, si c'est l'attitude à l'égard de l'objet dont il est question que l'on veut étudier. Mais on ne ferait que masquer le problème : celui-ci résulte de la complexité même de l'opinion, de ses éventuelles contradictions, du fait que, dans chaque cas particulier, différentes attitudes interagissent, et qu'il n'y a pas uniquement des maladresses d'expression de la part du rédacteur du questionnaire. Les connotations, positives ou négatives, des termes employés peuvent faire partie de la situation qu'on étudie. Il est plus important d'en prendre conscience, de les contrôler pour bien savoir ce que l'on fait, que de s'attacher à les éliminer complètement.

e) *Lorsqu'on propose aux sujets le choix entre plusieurs réponses, il est indispensable de s'assurer au préalable que la liste couvre bien toutes les positions possibles.* Les conséquences d'une omission peuvent aller très loin, et modifier l'ensemble des conclusions qu'on pourra tirer des réponses à la question.

Par exemple, dans le cas du questionnaire sur l'image du psychologue cité plus haut, si l'on n'avait pas proposé « le psychologue ne sert qu'à masquer les vrais problèmes de l'entreprise », non seulement on aurait empêché une des positions possibles de s'exprimer, mais en plus on aurait communiqué à l'éventuel lecteur des résultats l'impression globale que l'image du psychologue est dans l'ensemble favorable, et cela presque quelle que soit la répartition des réponses. Une telle impression serait inévitable, puisqu'une position franchement critique ou hostile n'aurait pas pu s'exprimer, si ce n'est par le refus de répondre, qui demande de la part du sujet une initiative rare, et qui risque d'être comptabilisé avec les « sans opinion ».

De tels oublis sont souvent l'expression des présupposés ou des préjugés du chercheur. Utilisés volontairement pour pouvoir ensuite présenter une image biaisée de l'opinion, ils constituent évidemment des fautes graves du point de vue déontologique. Mais, involontaires, ils sont tout aussi dangereux, scientifiquement cette fois : c'est une manière pour le chercheur de se fermer à ce qui ne va pas dans le sens de ses présupposés, et d'obtenir une pseudo-validation de ses hypothèses, qui le trompera lui-même.

Les présupposés du chercheur peuvent se glisser dans la formulation des questions elles-mêmes, même celle des questions ouvertes et pas seulement dans la gamme des réponses. Demander, par exemple, « Pour quelle raison avez-vous choisi votre métier ? » n'a de sens que si on admet qu'un métier est effectivement « choisi », et choisi pour des « raisons ». On impose au sujet un modèle de choix rationnel, où il apprécie les avantages et les inconvénients de chaque option, pour finalement choisir la meilleure. Cela exclut, implicitement, qu'on puisse se trouver avoir un certain métier comme aboutissement d'un processus tout différent. Si on laisse des personnes parler librement sur ce thème, ou si on les suit durant une période suffisamment longue de leur vie, on obtiendra souvent des descriptions où les choix et les raisons occupent peu de place ; les contraintes financières, la formation, la nature des informations disponibles, des éliminations pour des raisons contingentes, les occasions qui se présentent, des choix antérieurs qui n'ont eu sur le métier que des conséquences indirectes, etc., jouent souvent des rôles beaucoup plus décisifs. Pouvoir répondre à une question posée en termes de choix dépend à la fois de la situation de la personne interrogée dans la structure socio-économique (certaines personnes ont plus « le choix » que d'autres), et de la représentation qu'ils se font du processus en question.

Poser une question qui recouvre tout cela, sans exclure non plus, évidemment, les réponses en termes de choix rationnel, est très difficile. Si on connaît suffisamment les différents processus possibles, et si le contexte du questionnaire s'y prête, on pourra décomposer le problème, et poser le nombre voulu de questions précises. Si ce n'est pas possible, on pourra recourir à une question très large, du genre « Comment avez-vous été amené à pratiquer ce métier ? ». Certes, on peut reprocher à cette question d'être, elle aussi, biaisée, cette fois en sens inverse, puisqu'elle attire l'attention sur des événe-

ments : elle pourrait suggérer à la personne interrogée qu'elle a été uniquement l'objet d'un déterminisme extérieur. Mais comme cette représentation va à l'encontre des représentations les plus fréquentes, l'erreur ainsi introduite sera vraisemblablement faible.

Très souvent, les biais ainsi introduits sont très difficiles à déceler, parce que les présupposés du chercheur et ceux de ses sujets coïncident. Le questionnaire renvoie donc aux sujets, et aux futurs lecteurs des conclusions de l'enquête, la représentation qu'ils se font d'eux-mêmes et de leur situation, leur idéologie pour employer un langage différent. Il est donc possible que les personnes interrogées acceptent sans difficulté le cadre de pensée qui leur est proposé : c'est le leur, ou il n'en est pas éloigné. Si ce sont l'idéologie, les représentations, qu'on veut étudier comme telles, et si le chercheur a une conscience claire de cet objectif, alors la démarche peut être adaptée au problème, à condition qu'on se donne, comme dans toute démarche scientifique, la possibilité d'infirmer ses hypothèses. Mais si ce n'est pas le cas, si le chercheur prend les réponses pour une description de « la réalité » (qu'il s'agisse de faits ou d'opinions) l'enquête n'aura servi qu'à reproduire et à renforcer l'idéologie, en la présentant sous forme de résultat scientifique.

4.5.3.4. L'ORGANISATION DES QUESTIONS

Il arrive très souvent qu'on soit amené à poser plusieurs questions portant sur le même thème, ou sur des thèmes voisins, par exemple en vue de construire une échelle d'attitude, ou parce qu'on veut explorer les différents aspects du thème et analyser les relations entre les réponses. Il est alors nécessaire de mêler les énoncés favorables et défavorables. Certes, en principe, on devrait pouvoir se contenter de questions allant toutes dans le même sens, par exemple favorable : les sujets défavorables n'auraient qu'à répondre négativement. En fait, un tel questionnaire serait fortement biaisé. Il y a à cela plusieurs raisons, sur lesquelles il vaut la peine de s'attarder, car elles traduisent des mécanismes de portée générale.

La première est liée à l'existence, très souvent mise en évidence, d'un *biais d'acquiescement,* appelé aussi *biais de positivité* (en anglais *acquiescence set)* : toutes choses égales d'ailleurs, on a tendance à répondre « oui » plutôt que « non », « d'accord » plutôt que « pas d'accord ». Cette tendance est plus ou moins marquée selon les individus, et certains chercheurs y ont vu un trait de personnalité ; mais il semble que le contenu des questions joue aussi un rôle. Quelle que soit l'interprétation que l'on donne de ce biais, il faut reconnaître son existence, et essayer d'en minimiser la valeur, par exemple en alternant les questions où « oui » exprime une position favorable et celles où la même réponse traduit la position inverse. En équilibrant ainsi les énoncés favorables et défavorables, le biais est neutralisé, non pour chaque question, mais pour l'ensemble des questions qui portent sur le même thème.

La deuxième raison en faveur de l'alternance provient de la nécessité de neutraliser l'*effet de halo.* C'est aussi un phénomène bien connu mis en évidence en premier lieu dans l'analyse des jugements portés sur des élèves ou des travailleurs : on a constaté que lorsqu'on demande à une personne d'en juger une autre selon différents critères, l'ensemble des jugements est en général soit favorable, soit défavorable, comme si, en

fait, le juge se référait à une vue d'ensemble globale préalable, plutôt qu'à chaque critère séparément. Cet *effet de halo* est d'ailleurs renforcé lorsqu'on demande en premier lieu un jugement global : il semblerait que certaines personnes se sentiraient incohérentes si, ayant commencé par émettre un jugement d'ensemble favorable, elles exprimaient ensuite des critiques ou des réserves sur des points particuliers.

Ce même phénomène se retrouve dans les questionnaires, même lorsque l'objet à juger est abstrait. Une personne qui s'estime anti-raciste risque de donner à chaque question d'un questionnaire sur le racisme la réponse qu'elle juge la plus conforme à son attitude, sans se préoccuper vraiment du contenu spécifique des questions. L'alternance d'énoncés racistes et non racistes oblige à s'attarder, même pour peu de temps, sur le contenu de chaque question et de chaque réponse. Faute de quoi notre sujet anti-raciste, ayant réalisé que c'est la réponse « d'accord » qui exprime toujours sa position, la donnera à la limite sans avoir besoin de lire les questions.

La succession de plusieurs énoncés d'opinions convergentes peut agir comme une *source d'influence* susceptible de modifier les réponses ultérieures des sujets. Lire, ou entendre énoncer par l'enquêteur, une suite d'opinions allant toutes dans le même sens, revient à prendre connaissance d'un ensemble d'arguments en faveur de la position exprimée. Cela revient à présenter au sujet une argumentation partiale, unilatérale ; certes, il a la possibilité de la refuser, mais il faut pour cela qu'il trouve lui-même les arguments contraires, ce qui suppose un degré de réflexion préalable qu'on ne demande pas à celui dont l'opinion était déjà conforme aux énoncés proposés.

Enfin, lorsqu'on communique les résultats de l'enquête, le lecteur peut en retirer l'impression qu'une des positions n'a pas d'existence positive, mais ne serait que la négation ou le refus de la position opposée. Dans le domaine des attitudes politiques, par exemple, la droite ne se réduit pas simplement à l'opposition aux positions de la gauche : elle a ses valeurs, qui lui sont propres, et qui doivent donc pouvoir s'exprimer dans un questionnaire qui porte sur ce thème.

Pour qu'un questionnaire soit équilibré, proposer un nombre approximativement égal d'énoncés favorables et défavorables, n'est pas suffisant. Il faut encore qu'ils soient à peu près équivalents à d'autres points de vue. Par exemple, si l'on présente certaines opinions en les soutenant par une argumentation, alors que les autres sont présentées sans justification, on incite à approuver plutôt les premières. Par exemple, si l'on proposait les deux énoncés suivants, le choix entre eux serait certainement biaisé :

— L'avortement doit être autorisé, car il faut pouvoir n'avoir d'enfants que lorsqu'on le désire.

— L'avortement doit être interdit.

On suscite un biais de même type en présentant certains énoncés sous une forme extrême, et d'autres sous une forme plus modérée. Par exemple, toujours au sujet de l'avortement :

— L'avortement doit être autorisé sans aucune limitation.

— L'avortement doit être interdit, sauf dans un certain nombre de cas prévus par la loi.

Il est probable que certaines des personnes favorables à l'avortement refuseront la

première opinion, jugée trop radicale, et éventuellement accepteront la seconde, jugée par contraste plus acceptable.

Tout ce que nous avons dit à propos des questions où l'on demande aux sujets de choisir entre plusieurs opinions reste valable pour l'ensemble du questionnaire : il faut que les différentes positions qu'on peut prendre par rapport à l'objet étudie puissent s'exprimer. Nous ne reviendrons pas sur ce point, malgré son importance.

Ces diverses précautions portant sur l'équilibre d'un questionnaire ne témoignent pas seulement d'un souci déontologique d'impartialité. C'est aussi, et avant tout, un problème de validité scientifique de la recherche. Un questionnaire déséquilibré risque de tromper son auteur, avant même de tromper les lecteurs des conclusions.

4.5.3.5. Quelques thèmes particuliers

Tout ce que nous avons vu jusqu'à présent portait sur des exemples qu'on peut ranger, approximativement, dans le champ de l'opinion ou des attitudes. Nous allons examiner maintenant quelques problèmes particuliers qui, sans se distinguer radicalement des cas dont il a été question, peuvent faire appel à des techniques qui leur sont propres, ou soulèvent des difficultés plus spécifiques. Sans aucune prétention à l'exhaustivité, nous voulons passer en revue quelques exemples qui permettront de se rendre compte de la largeur et de la diversité des champs qu'on peut couvrir par la méthode des questionnaires, et aussi d'en préciser certaines limites.

a. *Les préférences*

Que ce soit pour mettre au point une campagne publicitaire, pour analyser les choix politiques, ou pour comprendre les goûts esthétiques, on peut être amené à étudier les préférences d'un ensemble d'individus.

A première vue, le problème est très simple : on dispose d'une collection d'objets (variétés d'un même produit, hommes politiques, œuvres d'art, etc.) et on veut savoir lequel est préféré ou, mieux, connaître l'ordre de préférence d'un groupe de personnes sur tous ces objets. Il semblerait qu'il n'y a qu'à le demander directement ; mais, déjà, à ce niveau, se pose le problème de choisir entre différentes techniques ; cette question peut être, en effet, posée sous différentes formes, dont les plus fréquemment utilisées sont :

a) Présenter tous les objets à la fois, et les faire ordonner totalement.

b) Présenter tous les objets à la fois, et les faire placer dans un nombre fixé de catégories ordonnées.

c) Présenter les objets un à un, et faire attribuer à chacun une note, par exemple entre 0 et 10.

d) Présenter successivement toutes les paires possibles d'objets, et faire indiquer dans chaque cas celui qui est préféré.

De nombreuses variations à partir de ces techniques sont possibles. On pourra, par exemple, constituer tous les sous-ensembles de n objets, et faire ordonner les objets dans chaque sous-ensemble (les comparaisons par paires en sont un cas particulier, où $n = 2$) ; lorsqu'on fixe le nombre de catégories dans lesquelles ranger les objets, on peut également imposer le nombre d'objets à mettre dans chacune ; lorsqu'on utilise les

comparaisons par paires, on peut postuler que les jugements sont certainement transitifs, ce qui évite de présenter certaines paires.

Ordonner tous les objets d'un ensemble, que ce soit totalement ou en les plaçant dans des catégories, est habituellement considéré par les sujets comme une tâche difficile ; c'est pourquoi on recourt volontiers à la méthode des comparaisons par paires, jugée plus simple. En revanche, elle peut être beaucoup plus longue, le nombre des paires possibles augmentant très vite avec le nombre d'objets ; si ce nombre est n, on constituera avec eux n(n - 1)/2 paires. Avec 10 objets, il faudra donc présenter 45 paires, ce qui est long et fastidieux, et les jugements deviennent de moins en moins fidèles à mesure que l'expérience se prolonge.

La technique qui consiste à faire attribuer des notes aux objets peut être considérée, d'un point de vue formel, comme équivalente à la méthode des catégories ordonnées, les notes définissant les catégories. En fait, du point de vue des sujets la tâche est très différente. Les deux consignes induisent des attitudes très différentes : mettre une note c'est le plus souvent porter un jugement absolu, comparer à un idéal, alors que placer des objets dans des catégories ordonnées incite plus à les comparer entre eux. Les deux techniques ne répondent donc pas exactement au même problème. En insistant sur l'aspect comparatif, on ne pourra pas, par exemple, mettre en évidence que tous les objets proposés sont insatisfaisants. En revanche, le fait de donner une note à chaque objet séparément peut amener à utiliser, dans chaque cas, des critères différents, ce qui pourra rendre les résultats difficilement interprétables.

Nous n'avons parlé jusqu'ici que de préférences, mais rien n'empêche de recourir aux mêmes techniques pour obtenir des jugements selon d'autres variables, tant que les jugements relatifs à l'ordre selon cette variable ont un sens. On pourra, par exemple, faire ordonner des professions selon leur prestige, des comportements selon qu'ils traduisent une attitude politique plus ou moins engagée, des objets selon leur commodité, etc. L'important est toujours de bien préciser aux personnes interrogées ce qu'on attend d'elles, quel critère on souhaite qu'elles utilisent. Pour prendre un exemple très simple, ce n'est pas la même tâche d'indiquer, devant deux boissons, laquelle on juge la plus sucrée, ou de dire laquelle on préfère du point de vue de sa teneur en sucre. L'analyse des résultats obtenus ne sera pas non plus la même : dans le premier cas, on pourra admettre que tous les jugements sont semblables, à des fluctuations aléatoires près, dans le second on n'a aucune raison de postuler cette homogénéité.

Les trois premières techniques nous donnent immédiatement l'ordre de préférence de chaque sujet, un ordre total dans le premier cas, un pré-ordre total dans les deux autres. Le premier problème qui se pose alors est de savoir comment traiter les écarts entre les ordres donnés par des individus différents, pour en tirer, éventuellement, un ordre « moyen » ou « représentatif » des préférences de la population, en supposant que cela ait un sens.

La quatrième technique, celle des comparaisons par paires, pose en plus un problème préalable : lorsque, comme cela arrive très fréquemment, certains des jugements ne sont pas transitifs (une personne peut dire préférer A à B, B à C, et néanmoins préférer C à A), est-ce que cela a encore un sens de parler d'un ordre de préférences du sujet en question ? Si oui, comment déterminera-t-on cet ordre ?

Ces problèmes ont donné lieu à de très nombreux travaux et à une abondante littérature. Les solutions proposées varient selon les postulats admis quant aux mécanismes de comparaison et de choix, quant à la nature probabiliste ou non des jugements, et quant à l'homogénéité de la population. Nous y reviendrons plus loin, dans la partie consacrée aux techniques d'analyse. Mais il nous faut dès maintenant en indiquer certaines conclusions, car elles conditionnent la manière de poser les problèmes, et déterminent ce qu'on peut attendre de chaque technique.

Il faut réaliser que la question des « préférences d'une population » n'a pas nécessairement une réponse. Si toutes les personnes interrogées sont unanimes, si elles ont toutes donné le même ordre, alors évidemment il n'y a pas de problème : cet ordre représentera adéquatement les préférences du groupe. Mais ce cas simple est très rare.

Quand on ne s'intéresse qu'à deux objets, on peut éventuellement se contenter de voir quel est le jugement majoritaire, ou indiquer quelle est la proportion de chacun des deux jugements possibles. Mais dès qu'on passe à trois objets, le problème se complique considérablement. Avec 3 ! = 6 ordres possibles, la notion d'ordre majoritaire peut ne plus avoir beaucoup de sens, à moins que l'un d'eux ne soit très nettement plus fréquent que tous les autres.

On résoudra souvent le problème en ne considérant pas les préférences comme simplement ordinales, c'est-à-dire en traitant les rangs comme des valeurs dont on peut prendre la moyenne, ce qui suppose au moins que les écarts entre deux rangs successifs sont égaux. On ordonne ensuite les objets selon leur rang moyen. Cette opération est formellement toujours possible. Mais on peut se demander quel sens a l'ordre obtenu, lorsque les préférences des individus sont trop divergentes. Dans un exemple où les six ordres possibles seraient également fréquents, les trois objets devraient être considérés comme équivalents, chacun ayant le même rang moyen. Mais on commettrait évidemment une très grave erreur en supposant que tous les membres de la population sont indifférents entre ces trois objets. Dans des cas comme celui-ci, il vaut mieux reconnaître que la population n'est pas homogène dans ses préférences, et donc que parler des « préférences de la population » n'a aucun sens. A la rigueur, on peut énoncer, sur la population, une proposition négative : il n'y a pas d'objet qui, majoritairement, soit préféré aux autres [1].

Plutôt que de chercher ainsi à décrire les préférences de la population, on peut se demander quels sont les déterminants des préférences individuelles ; par exemple, on se demandera selon quelles autres variables ceux qui ont donné chacun des ordres se distinguent des autres. C'est tout à fait possible lorsque les objets sont en petit nombre : avec 3 objets on peut avoir au maximum 3 ! = 6 ordres de préférence ; on constitue ainsi au plus 6 catégories de sujets dont il n'est pas difficile, pratiquement, d'étudier les caractéristiques. Mais il suffit qu'on ajoute un objet, et on passe à 4 ! = 24 ordres possibles, donc à 24 catégories. Même si toutes ne sont pas représentées, c'est une variable difficile à manipuler avec des échantillons de taille courante.

---

1. Le coefficient W, dit *coefficient de concordance*, permet de vérifier l'hypothèse que tous les groupes sont équiprobables, donc qu'il est inutile de chercher un ordre représentatif. Si cette hypothèse est rejetée, l'ordre des rangs moyens peut être considéré, avec prudence, comme une sorte de « rang moyen ». Si l'on ne peut pas ainsi résumer l'ordre de préférence dans une population, on pourra découper celle-ci en groupes supposés plus homogènes, et vérifier si, à l'intérieur de chacun d'eux, on observe un accord supérieur.

Il faut alors se résigner à n'utiliser qu'une partie de l'information contenue dans les observations ; par exemple, on pourra caractériser un individu non par tout son ordre de préférences, mais par l'objet qu'il a placé en tête, ou les deux objets placés en tête, ou par le premier et le dernier. Mais il faut toujours prendre garde au fait que lorsqu'on construit une variable par une opération combinatoire, le nombre de modalités de cette variable croît très vite, et, ce qui est la même chose, le nombre des groupes qu'elle permet de distinguer.

C'est au moment de la construction du questionnaire qu'il faut se poser ces questions. Si les analyses prévues ne permettent pas de prendre en considération l'ordre total exprimé par chaque sujet, il est probablement inutile de le lui demander ; une consigne plus simple, qui se limitera à l'information utilisable, peut éventuellement suffire.

Jusqu'à présent nous avons considéré les objets comme donnés. Or, effectivement, dans les exemples que nous avons cités, c'était le cas, et l'ordre de préférence sur ces objets nous intéressait directement. Mais on peut aussi se poser d'autres problèmes, par exemple se demander en fonction de quels critères se font les choix, et comment ces critères sont combinés pour permettre un jugement global de préférence sur les objets. De telles questions amènent à *choisir* les objets présentés, non selon leur intérêt intrinsèque, mais en fonction d'un plan expérimental permettant, par exemple, de voir ce que donne la comparaison de deux objets qui ne diffèrent que selon un seul critère. Mais cela suppose que les objets correspondent à toutes les combinaisons des différents critères considérés, ce qui n'est pas toujours le cas. S'il est possible de présenter des aliments qui varient systématiquement selon le goût, la consistance, la couleur, etc., c'est évidemment beaucoup plus difficile pour, par exemple, des hommes politiques ou, dans un domaine différent, pour des logements. On en est alors réduit à prendre non pas les objets eux-mêmes, mais leur *description*, au moyen des critères retenus. On présentera non plus des logements réels, mais simplement des logements décrits par leur nombre de pièces, leur prix, leur environnement, etc. Cela suppose d'abord qu'on a des raisons suffisantes d'admettre qu'on connaît les critères importants, et qu'on peut sans inconvénients graves les isoler des autres. Ensuite, cette méthode, en attirant l'attention sur les critères, ne laisse possible qu'une démarche analytique, probablement très différente de l'appréhension plus globale des objets réels. Néanmoins, lorsque les objets n'existent pas encore, ou lorsqu'ils sont trop vastes pour être matériellement présentés au cours de l'enquête, ou trop insuffisamment connus pour être simplement désignés par leur nom, on est obligé de recourir à ces descriptions, tout en étant conscient de leurs limites. Enfin, nous retrouvons une difficulté technique déjà rencontrée : lorsque le nombre des critères retenus, ainsi que celui des modalités de chaque critère, augmentent, le nombre de combinaisons à présenter croît très vite, et rend les jugements pratiquement impossibles. On est donc réduit à ne prendre en considération qu'un nombre faible de critères simultanément.

Devant ces difficultés, et puisque, de toute manière, la tâche demandée aux sujets est artificielle, on la simplifie parfois en leur proposant de comparer non des objets, mais directement des critères. On demandera, par exemple, si, pour un certain produit, le prix est plus ou moins important que la qualité (en supposant qu'on ait défini sans équivoque ce qu'on entend par ce dernier terme). C'est une consigne qui est en général

bien acceptée, et la plupart des sujets répondent facilement à des questions de ce genre. Malheureusement, lorsqu'on les examine de près, les réponses sont très ambiguës, et les conclusions qu'on peut en tirer peu utilisables.

En effet, dire que la qualité est plus importante que le prix ne signifie pas que, dans tous les cas, on choisira le produit de meilleure qualité, quel que soit le prix. Il y a un prix limite, plus ou moins bien défini, qu'on n'acceptera pas de dépasser. La question qui a un sens, c'est : « Quel est l'accroissement maximum de prix que vous êtes disposé à accepter pour une amélioration définie de la qualité ? ». En toute rigueur, il faudrait aussi poser la même question pour une diminution du prix et de la qualité, la compensation entre les deux pouvant ne pas se faire de la même manière dans les deux sens. Enfin, il faudrait expliciter le cadre de référence, qui est probablement constitué par le prix et la qualité actuels. C'est probablement tout cela qui est sous-jacent aux réponses des sujets, lorsqu'on leur demande de comparer directement des critères, et il se peut que des questions plus précises et moins ambiguës, mais plus artificielles ne fassent que les troubler. Avant de choisir la technique, il faudrait, au cours de tests préalables, essayer de faire expliciter, dans chaque cas particulier, l'interprétation des questions par les sujets, et la signification qu'ils donnent à leurs réponses.

Faire comparer directement des critères, c'est aussi, implicitement, supposer qu'il n'y a pas d'interaction entre eux. Par exemple, on a constaté que lorsqu'on demande de juger ainsi des modes de transport, le fait d'être assis est déclaré plus important que le prix pour des trajets longs, mais pas pour des trajets courts. Une décomposition critère par critère, qui ne tiendrait pas compte des effets d'ensemble, risque de donner des résultats très artificiels.

b. *La satisfaction*

Evaluer, « mesurer » comme on le dit souvent, la satisfaction d'un groupe de personnes par rapport à un certain objet ou une situation, est une demande souvent adressée aux spécialistes des enquêtes. Que ce soit pour évaluer a posteriori la réussite ou l'échec d'une action, pour effectuer des prévisions, ou dans des buts plus immédiatement politiques, comme dans le cas des comparaisons internationales, les enquêtes de satisfaction se multiplient. On interroge des personnes sur leur satisfaction générale dans leur vie, sur leur satisfaction par rapport à leur logement, leur environnement, leur travail, leur famille, etc.

Les questions posées sont, en général, étonnamment simples. Néanmoins, la cohérence des réponses obtenues avec d'autres données, ainsi que, dans certains cas, leur valeur prédictive, indiquent qu'elles ont une validité certaine. On a utilisé, par exemple :

« Dans l'ensemble, estimez-vous que vous êtes
— très satisfait
— plutôt satisfait
— plutôt insatisfait
— très insatisfait
de votre vie actuelle ? »

Parfois, on a remplacé « satisfait » par « heureux », « insatisfait » par « mécontent ».

On peut poser le même type de questions par rapport à des objets particuliers, tels que le logement, le travail, etc. On constate, par exemple, une bonne corrélation entre la note moyenne de satisfaction au travail dans une entreprise et le *turn-over* dans cette même entreprise. Poser, au cours de la même enquête, des questions portant à la fois sur la satisfaction globale et sur la satisfaction relative à des domaines particuliers, peut éventuellement permettre d'estimer le poids de chaque domaine dans l'impression générale de satisfaction. Mais on se heurte facilement ici à l'effet de halo, ou de cohérence : les personnes qui se déclarent globalement satisfaites ont tendance à se dire également satisfaites sur des points particuliers.

Du point de vue théorique, de telles mesures de satisfaction peuvent évidemment être considérées comme grossières. Peut-être ne devrait-on pas accorder à chaque jugement une valeur absolue, mais l'interpréter en le rapportant au niveau d'aspiration ou aux expériences passées des sujets. H. Cantril, par exemple, a utilisé, notamment pour des comparaisons internationales, un type particulier d'échelle, dites « d'auto-ancrage » *(self-anchoring scales).* On présente à la personne interrogée une échelle en 13 points, en expliquant que la valeur minimum représente la vie la plus malheureuse qu'on puisse imaginer. On demande au sujet de décrire cette vie ; puis on passe à l'autre extrémité de l'échelle, qui représente la vie la plus heureuse. Le sujet ayant à nouveau décrit cette vie, on lui demande alors de situer sur l'échelle ainsi définie sa vie actuelle, éventuellement sa vie passée ou sa vie telle qu'il l'imagine dans quelques années.

Cette manière un peu complexe de procéder présente deux avantages. En premier lieu, c'est le sujet qui, en quelque sorte, construit lui-même son échelle : il est très clair que c'est par rapport à ses propres conceptions de la vie heureuse ou malheureuse qu'on lui demande d'apprécier sa vie actuelle. Deuxièmement, les descriptions des extrémités de l'échelle fournissent des éléments intéressants en eux-mêmes sur le système de valeurs des personnes interrogées, leurs craintes, leurs aspirations.

L'idée sous-jacente est évidemment que la satisfaction ne peut être appréciée que subjectivement, et par rapport aux normes propres de chaque personne interrogée. L'explicitation du contenu des extrémités a pour objet de renforcer cet aspect personnel, plus que si l'on parle en termes généraux et abstraits de satisfaction ou de bonheur, et sans courir le risque, en précisant, d'imposer un point de vue extérieur. Des personnes pourront donc être considérées comme ayant la même satisfaction, même si le contenu du « malheur absolu » de l'une n'est pas très éloigné du « bonheur parfait » de l'autre.

c. *Intentions et anticipations*

Il est très fréquent qu'on attende du psychologue ou du sociologue qu'il prévoie les comportements de certaines personnes dans des situations non encore réalisées. Un produit nouveau se vendra-t-il ? Une mesure sociale aura-t-elle les effets souhaités ? Que se passera-t-il si on transforme la structure d'un quartier ? Quelles seront les conséquences de l'atténuation des tabous sexuels ? Les spécialistes des sciences sociales

sont considérés comme connaissant les « besoins », et donc sont censés pouvoir répondre à de telles questions.

En supposant qu'ils soient effectivement en mesure de le faire, ce qui n'est pas du tout évident, et n'est certainement pas vrai dans tous les cas, il faut se demander si c'est bien l'enquête qui est l'instrument le mieux adapté à cette catégorie de problèmes. Certes, on peut toujours poser des questions directes du genre « Que feriez-vous si... ? », on obtiendra vraisemblablement des réponses, et même des réponses cohérentes. Mais on peut douter de leur valeur prédictive, et ceci pour plusieurs raisons.

Tout d'abord, la description verbale qu'on peut donner d'une situation nouvelle, future, ne peut pas l'épuiser ; il reste toujours des aspects, parfois difficiles à formuler, auxquels ne seront peut-être sensibles que certaines personnes et que la description omettra. D'autre part, la description induit chez les sujets interrogés une attitude analytique, où chaque élément de l'objet ou de la situation risque d'être jugé séparément des autres, alors que devant la situation réelle on aura une perception et des réactions plus globales. Enfin, ce type de question ignore les aspects temporels des réactions à des situations nouvelles : inertie des habitudes antérieures, essais, apprentissage ou invention de conduites nouvelles, imitation d'autres personnes, constitution progressive d'une image sociale, etc. D'autre part, on s'est aperçu que certaines personnes ne répondaient pas pour elles-mêmes, mais pour les autres. Par exemple, lorsqu'on demande, au cours d'une enquête, si on souhaite une piscine dans un quartier, la réponse est presque toujours affirmative, même si l'on n'a aucune intention de la fréquenter soi-même.

Reconnaissant l'existence de telles limites, il serait dangereux de se fonder sur des enquêtes de ce genre. L'idéal, pour répondre au problème posé, serait de faire une expérience, si possible en vraie grandeur. Certaines entreprises, avant de lancer un produit nouveau à grande échelle, le diffusent dans une ville ou dans une région, et y observent les réactions des consommateurs. Avant d'imposer la limitation de vitesse sur toutes les routes, elle a été essayée sur quelques-unes seulement, à titre expérimental. Des modifications complètes de la circulation ont été tentées, pendant des périodes brèves, dans le centre de certaines villes. A condition que les modalités soient bien analysées, qu'on dispose de termes de comparaison et de critères de jugement adéquats, et que la durée de l'expérimentation soit suffisante pour saisir tous les aspects temporels, de telles expériences en vraie grandeur constituent, évidemment, la méthode la meilleure. Dans ces méthodes, l'enquête retrouve un rôle, non pour évaluer des préférences ou prévoir des comportements, mais pour obtenir des informations plus fines que les données globales que l'on peut recueillir (ventes, nombre d'accidents, etc.). En permettant une analyse plus approfondie de la manière dont la nouvelle situation a été perçue, des opinions et des jugements exprimés à son propos, on pourra mieux comprendre les processus d'adaptation ou de rejet, prendre des mesures mieux adaptées et faire de meilleures prévisions quant à ce qui pourra se passer sur d'autres terrains.

Il est évident toutefois que de telles expériences en vraie grandeur ne sont pas toujours possibles : elles peuvent être coûteuses, politiquement peu acceptables, elles peuvent déclencher des processus qui rendent le retour en arrière impossible, même en cas

d'échec ; celui-ci peut avoir des conséquences trop graves pour qu'on veuille en courir le risque sans indications préalables, les conclusions sur l'expérience ne pouvant être tirées que trop tard, etc. Pour toutes ces raisons, il est fréquent que la demande d'enquête se maintienne, soit comme substitut, soit comme préalable à une telle expérimentation.

Dans ce cas, plutôt que de poser des questions sur des intentions ou des prévisions de comportement dans des situations hypothétiques, il est préférable d'analyser de façon aussi approfondie que possible la situation présente, afin de mettre en évidence les critères de jugement utilisés, les insuffisances ressenties, les besoins exprimés. Toutes ces informations étant obtenues par rapport à la situation actuelle, elles ont des chances d'être plus réalistes que les réactions par rapport à des situations hypothétiques. Celles-ci ne sont pas d'ailleurs à éviter absolument : elles peuvent, dans certains cas, constituer des révélateurs d'attitudes intéressantes à connaître. Simplement, il ne faut pas directement fonder les prévisions sur elles.

4.6. LES PRÉ-TESTS DU QUESTIONNAIRE

Lorsqu'une première version du questionnaire est rédigée, c'est-à-dire lorsque la formulation de toutes les questions et l'ordre de celles-ci sont fixés, à titre provisoire, il est nécessaire de s'assurer que le questionnaire est bien applicable, et qu'il répond effectivement aux problèmes que se pose le chercheur. Pour cela, une série de vérifications empiriques sont nécessaires. C'est l'ensemble de ces vérifications qui constituent les pré-tests, où il ne faut pas voir simplement un essai du questionnaire à petite échelle, qui ne représente que l'une des opérations possibles. Selon les problèmes qui se posent, les points dont on veut s'assurer, on sera amené à recourir à différentes techniques.

Lorsqu'un questionnaire est rédigé, on peut s'interroger sur les points suivants :

— Toutes les questions sont-elles comprises, sont-elles comprises de la même manière par tous, et de la manière prévue par le chercheur ?

— Certaines questions ne paraissent-elles pas trop difficiles ?

— Les listes de réponses proposées aux questions fermées recouvrent-elles toutes les réponses possibles ?

— Toutes les réponses sont-elles acceptées par les sujets ? N'y en a-t-il pas qui susciteraient trop de refus inutilisables ?

— L'ordre des questions est-il acceptable ? N'y a-t-il pas trop de ruptures, de coq-à-l'âne ? Certaines questions ne risquent-elles pas d'influencer les réponses aux questions suivantes ?

— N'y a-t-il pas des questions inutiles, soit parce qu'on manquera d'informations complémentaires pour en interpréter les réponses, soit parce que la quasi-totalité des sujets donneront la même réponse ?

— Comment les sujets réagissent-ils à l'ensemble du questionnaire ? Ne le trouvent-ils pas trop long, ennuyeux, difficile, indiscret, partial ?

En dehors du questionnaire lui-même, on peut se poser des questions sur la manière

de choisir l'échantillon, de prendre contact avec les futurs sujets, sur les conditions de passation, etc.

Grossièrement, il faut distinguer deux phases dans le pré-test d'un questionnaire. La première porte sur chaque question considérée en elle-même, la seconde sur le questionnaire dans sa totalité et sur les conditions de passation.

La première phase pourra s'effectuer auprès d'un petit nombre de personnes (parfois une dizaine peut suffire), en choisissant de préférence des sujets appartenant à des milieux suffisamment différents de celui des auteurs du questionnaire. On leur pose les questions telles qu'elles sont formulées, mais en leur demandant des réponses développées, commentées, et des remarques sur la signification qu'elles attribuent à la question. A la limite, ce pourrait être un petit entretien non directif dont la question serait la consigne.

Lorsqu'il est prévu de fermer une question, on pourra, au cours de cette phase, commencer par la poser sous forme ouverte, enregistrer la réponse complète, puis présenter la liste des réponses prévues, en demandant d'en choisir une. Cela permettra d'apprécier la correspondance entre les deux types de réponses, de savoir quels discours correspondent à chacune des réponses proposées. On se rendra aussi compte si la liste est insuffisante, ou si elle suggère des réponses auxquelles le sujet n'avait pas pensé spontanément. Lorsqu'on hésite entre deux formes d'une même question, un pré-test de ce type peut fournir des éléments de choix.

Bien qu'une des caractéristiques de cette phase soit d'obtenir des sujets des réponses beaucoup plus détaillées que lors de la passation finale, il faut veiller à ne pas trop les « pousser ». En les incitant à réfléchir aux questions, à commenter leurs réponses, à verbaliser leurs réactions, on risque d'induire chez eux une attitude très différente de celle d'une personne à qui on demande de répondre à un questionnaire fermé, et ainsi de l'amener à un niveau de compréhension, de réflexion et de critique très nettement supérieur.

Cette première phase du pré-test du questionnaire nous indique comment les questions et les réponses sont comprises, elle nous permet d'éviter des erreurs de vocabulaire ou de formulation, de mettre en évidence des refus, des incompréhensions et des équivoques. Mais cela ne nous assure ni de l'acceptabilité du questionnaire dans son ensemble, ni de sa bonne adaptation aux besoins de la recherche.

Pour vérifier ce dernier point, avant d'entreprendre des investigations empiriques, on peut commencer par faire le plan de dépouillement complet, c'est-à-dire prévoir dans le détail toutes les opérations qu'on effectuera sur les réponses, une fois recueillies. Mieux, on peut se donner la peine d'écrire tous les tableaux des résultats qu'on désire obtenir, en notant dans chaque cas des valeurs fictives, mais vraisemblables. Cet exercice, bien que parfois très ennuyeux, permet de s'apercevoir de certains oublis [1]. On constatera, par exemple, que les réponses à telle ou telle question sont impossibles à interpréter si on ne dispose pas, par ailleurs, de renseignements supplémentaires qu'on n'avait pas pensé recueillir, parce qu'ils ne présentaient pas d'intérêt en eux-mêmes. Construire des tableaux fictifs amène aussi à se demander comment on interprétera des

1. C'est également lui qui permet de déterminer la taille de l'échantillon nécessaire (cf. 2.2.6.).

résultats non conformes aux hypothèses, ce qui peut en suggérer de nouvelles qui, à leur tour, amènent à poser d'autres questions. On pourra aussi s'apercevoir que certaines questions sont inutiles, soit parce qu'elles en recoupent d'autres, soit que l'information qu'elles apportent est sans intérêt.

Lorsqu'on a ainsi anticipé sur le dépouillement et modifié le questionnaire s'il y a lieu, lorsqu'on s'est assuré que chaque question est acceptable, on peut passer à un essai à petite échelle, dans des conditions si possible identiques à celles de la passation définitive. Il ne s'agit plus maintenant de vérifier seulement quelques points précis, c'est à la fois le questionnaire, la procédure d'échantillonnage et l'équipe d'enquêteurs qu'on met à l'épreuve.

Lorsqu'on en a les moyens, il faudrait pouvoir interroger ainsi une cinquantaine de personnes au moins. On pourra ainsi évaluer le taux de refus, se rendre compte de la manière dont les sujets réagissent au questionnaire, si l'ordre des questions ne pose pas de problème.

Une première analyse des résultats obtenus donnera la possibilité d'évaluer, ne serait-ce que grossièrement, la valeur de certaines hypothèses, ou d'esquisser certaines interprétations. Pendant que l'enquête définitive se déroule, on pourra gagner du temps en analysant les réponses aux questions ouvertes pour en préparer le codage, ou vérifier si les réponses aux questions d'attitude permettent bien de construire les échelles prévues.

On pourra aussi constater qu'il y a des questions auxquelles presque tous les sujets répondent de la même manière, ce qui les rend très peu utiles pour des analyses plus fines, par croisement avec d'autres questions. Ce peut être dû soit à un accord réel, auquel cas la question est inutile, ou à une maladresse de formulation, qui oriente les sujets vers une des réponses. Il faut alors rectifier la manière dont la question est posée.

A la fin de cette pré-enquête, il est très utile, pour le chercheur, de s'entretenir avec les enquêteurs, qui peuvent signaler les difficultés qu'ils ont rencontrées, les types de refus qu'on leur a opposés, les incompréhensions qu'ils auront cru déceler chez les sujets, les réactions d'ennui ou d'impatience, etc. On court toutefois le danger, en poussant trop ces discussions, de s'arrêter plus qu'il ne le faudrait à des anecdotes, à des cas isolés.

Les pré-tests que nous venons de décrire constituent une phase longue et coûteuse, souvent incompatible avec le budget total de l'enquête. Il faut parfois abréger ou supprimer telle phase, se contenter d'indications sommaires sur certains points. C'est un problème d'utilisation optimum des ressources dont on dispose. Mais il faut bien tenir compte du fait que, lorsque le travail définitif sur le terrain est commencé, on est entré dans une phase irréversible. Même si on se rend compte de certaines erreurs, il n'est très souvent plus possible de les corriger, sauf à tout recommencer.

Chapitre 5

LE DÉROULEMENT D'UNE ENQUÊTE

5.1. LES CONDITIONS DE PASSATION

5.1.1. Les lieux

Le choix de la technique et les conditions dans lesquelles se fera l'enquête ne sont pas indépendants. L'entretien non directif est incompatible avec les contacts dans la rue ou dans d'autres lieux où les personnes à interroger ne font que passer. On ne construira pas un questionnaire de la même manière selon que les questions seront posés par l'enquêteur, qui notera les réponses, ou qu'on laissera un formulaire au sujet, qui le remplira lui-même.

Toutefois les avis convergent pour estimer que, quelle que soit la technique adoptée, c'est l'enquête au domicile du sujet qui assure les conditions les plus favorables. On y dispose, en général, d'un temps suffisant, les personnes interrogées se sentent à l'aise, on peut réduire au minimum les risques d'être dérangé ou écouté. Le seul inconvénient notable réside dans la difficulté où l'on se trouve souvent de voir l'enquêté seul ; il est parfois délicat d'expliquer à une personne qu'on veut la voir seule, et de demander à son conjoint de quitter la pièce, ce qui peut d'ailleurs être difficile dans certains logements exigus. Pourtant il est très probable que la présence d'une tierce personne, surtout un proche, modifie certaines réponses.

Les enquêtes sur les lieux de travail posent beaucoup de problèmes. Dans la mesure où l'enquêteur doit en général avoir obtenu l'autorisation de la direction, il peut être considéré comme son employé. Les risques d'être dérangé, d'avoir des collègues qui écoutent ce qui se dit, sont importants bien que parfois on puisse les diminuer en faisant passer l'enquête dans une pièce ad hoc ; mais cela renforce l'impression que l'enquête est faite pour la direction. En interrogeant des personnes sur leur lieu de travail, c'est dans leur rôle de travailleur qu'elles se situent. Il faut donc, dans la mesure du possible, ne le faire que lorsque le thème de l'enquête est effectivement lié à l'activité professionnelle ou aux conditions de travail.

Partout ailleurs, il faut se contenter de questionnaires brefs, qui ne devraient pas dépasser les dix minutes, et même être nettement plus courts si l'on n'a pas la possibilité de s'asseoir.

Les lieux publics où il arrive qu'on reste assez longtemps sans activité précise, comme les cafés ou les squares, sont assez favorables : les refus n'y sont pas trop nombreux, mais il est impossible d'obtenir l'isolement souhaitable. D'autres personnes passent, s'approchent, commentent ce qui se dit, il y a souvent un bruit de fond important qui peut devenir très gênant lorsque l'entretien est enregistré au magnétophone. Enfin, il est évident que les populations qu'on peut y toucher sont très particulières.

Les lieux où l'on attend, comme les salles d'attente de gare ou les couloirs des administrations sont de faux bons endroits pour effectuer des enquêtes. Certes, les personnes y sont disponibles, mais, outre que l'isolement y est impossible, les sujets sont facilement tendus, craignent de laisser passer leur tour. De plus, accepter de répondre c'est reconnaître qu'on attendra encore longtemps, ce qui est désagréable. De façon générale, les lieux où l'on est anxieux sont très peu propices à des entretiens.

Lorsque la nature du problème posé exige un échantillonnage sur place (cf. 2.2.2.), et en même temps un entretien ou un questionnaire de longue durée, on peut se contenter d'un premier contact, en demandant simplement un rendez-vous au domicile de la personne qu'on veut interroger. Le seul inconvénient est que certains se méfient de cet inconnu qui demande leur adresse, et refusent, ou même, cela est arrivé, le signalent à la police.

5.1.2. Le contact

Un inconnu qui accoste dans la rue, ou qui sonne à une porte, a de fortes chances de rencontrer une réaction de méfiance. Il lui faut donc immédiatement indiquer pourquoi il est là, et ce qu'il attend. Il est préférable, dans ce contexte, d'employer le mot « sondage », qui commence à être connu du grand public, de préférence à « enquête » dont les connotations policières peuvent rendre méfiant.

Lorsqu'on dispose d'une base de sondage indiquant les noms et les adresses des personnes à interroger, annoncer l'arrivée de l'enquêteur par une lettre facilite considérablement le premier contact. Cette lettre devrait comporter le nom et l'adresse de l'organisme enquêteur et des indications sur le but de l'enquête, éventuellement l'ordre de grandeur de la durée de l'entretien. On peut ajouter la garantie de l'anonymat des

réponses, et assurer qu'on n'a rien à vendre. Si la personne a le téléphone, on peut l'utiliser pour prendre rendez-vous, mais cela ne concerne évidemment qu'une population encore restreinte.

Lorsque la base de sondage manque, et qu'il faut se rabattre sur les entretiens dans la rue ou sur le porte-à-porte, le contact est évidemment plus difficile, et il faut pouvoir donner toutes les indications nécessaires avant que la porte ne se referme. Bien qu'on n'ait guère réussi, jusqu'à présent, à donner des indications systématiques à ce sujet, il est certain qu'il y a des enquêteurs qui sont mieux acceptés que d'autres.

Il est difficile de déterminer ce qu'on va dire au futur sujet du but de l'enquête et des thèmes qui seront abordés. La déontologie exigerait que les personnes contactées puissent accepter ou refuser en pleine connaissance de cause. Mais c'est parfois difficile, pour des raisons diverses. Tout d'abord, certains clients tiennent à ce qu'on ne sache pas qu'ils ont commandé une enquête sur un thème donné. Qu'il s'agisse d'entreprises industrielles qui craignent de fournir à leurs concurrents des éléments d'information sur leur stratégie commerciale, ou d'administrations qui craignent des retombées politiques, des conditions de secret de ce type sont fréquentes.

Des raisons proprement scientifiques peuvent aussi s'opposer à ce qu'on indique d'emblée aux sujets pour qui on travaille, et même l'objectif réel de l'enquête. Il arrive en effet souvent, et nous l'avons vu en particulier à propos des consignes des entretiens non directifs, qu'on choisisse d'aborder pour commencer des thèmes très larges, pour pouvoir, ensuite, situer dans ce contexte ce qui sera dit sur des points plus précis, et aussi pour savoir si ces points sont abordés spontanément. Il est évident qu'en annonçant d'emblée l'objectif final, on court-circuite les étapes prévues et qu'on rend inapplicable la stratégie qu'on avait adoptée. De même, en indiquant clairement au sujet pour qui on travaille, on risque de l'inciter à tenir un discours qui s'adresse spécifiquement au client, sous la forme, selon le cas, d'absence de critiques, ou de revendications systématiques. Ces difficultés ne sont pas propres aux entretiens, on les retrouve identiques pour les questionnaires.

Quelle que soit la solution adoptée par le chercheur face à ces exigences contradictoires, il est important qu'elle se traduise par des consignes très claires données aux enquêteurs, pour assurer la constance des conditions de passation. L'enquêteur doit savoir exactement ce qu'il doit dire, les réponses qu'il peut donner aux éventuelles questions, la manière d'éluder celles-ci si c'est nécessaire, ce qu'il peut dire après la fin de l'entretien. Une manière d'éviter certaines questions peut en effet consister à dire qu'on y répondra plus tard, et il faut évidemment le faire. Notons toutefois que peu de personnes, même lorsqu'elles acceptent de se prêter à l'enquête, manifestent une grande curiosité.

Les rôles respectifs de l'enquêteur et de l'enquêté doivent être très clairement définis par l'enquêteur : s'il s'agit d'un entretien non directif, il faut que le sujet sache qu'il ne lui sera pas posé de questions. Le questionnaire demande moins de précisions préalables, car il correspond plus à ce quoi s'attendent la presque totalité des sujets.

Lorsque l'échantillonnage est fait par la méthode des quota, l'enquêteur doit, avant de commencer l'entretien ou le questionnaire proprement dit, s'assurer que la personne qu'il a en face de lui appartient bien à l'une des catégories qu'il doit interroger. Il est

quelquefois désagréable, surtout pour l'enquêteur qui vient d'essuyer plusieurs refus et qui rencontre enfin une personne de bonne volonté, de lui expliquer qu'il ne peut pas l'interroger.

5.1.3. Le déroulement

Il est fixé par tout ce que nous avons dit sur les entretiens et les questionnaires : la consigne et les règles de l'entretien non directif, le guide de l'entretien centré, le questionnaire, doivent indiquer en principe à l'enquêteur *tout* ce qu'il a à faire. Les consignes qu'on lui donne doivent aussi porter sur des cas relativement rares, comme celui où la personne interrogée demande des explications, où elle n'a manifestement pas compris une question, où elle se met à poser à son tour des questions à l'enquêteur sur ses propres opinions ; lorsqu'il s'agit d'une question de fait, il faut savoir si l'on accepte ou non que le sujet aide sa mémoire par le recours à un document quelconque, un agenda par exemple, ou demande la réponse à une autre personne.

Cette insistance sur la rigueur des consignes aux enquêteurs peut sembler exagérément rigide, et induire chez eux un comportement artificiel inadapté à certaines situations. Ne serait-il pas préférable, dans certains cas, de laisser l'enquêteur juger de la réaction nécessaire, et de la conduite la plus adéquate à la personne en face de lui ? L'information recueillie n'en sera-t-elle pas plus riche ? Il est vrai que toute règle à ses limites, et qu'il y a des cas où, certainement, s'en écarter serait bénéfique. Si un sujet a manifestement mal compris une question, il ne sert à rien de faire semblant de ne pas s'en apercevoir, de la noter et de poursuivre comme si de rien n'était. Mais de tels cas, dans l'ensemble, sont rares, et il est préférable de courir ce risque que le risque opposé, qui serait de rendre chaque entretien ou questionnaire différent des autres, de supprimer toute comparabilité, par la fluctuation des conditions de passation.

Les entretiens non directifs ou centrés sont presque toujours enregistrés au magnétophone, pour des raisons évidentes : c'est le seul moyen de conserver intégralement ce qui a été dit, aussi bien par le sujet que par l'enquêteur, et celui-ci, déchargé du souci de prendre des notes, peut consacrer toute son attention à ce qui est dit et au contrôle de ses propres réactions. Parfois, on emploie aussi le magnétophone pour des questionnaires très brefs qu'on fait passer dans la rue ou dans un lieu public, où les sujets sont pressés. L'enregistrement permet de gagner du temps sur la transcription des réponses, même si elle se ramène à mettre des croix dans des cases. Mais les bruits environnants prennent une grande importance à l'écoute, ce qui peut rendre les réponses inaudibles. Ce risque existe toujours en plein air ou dans des lieux publics, mais des micros directionnels peuvent l'atténuer.

Néanmoins, le principal inconvénient du magnétophone réside dans l'effet inhibiteur qu'il peut avoir sur la personne interrogée. Certains le refusent : ce sont surtout des personnes interrogées ès qualité, hommes politiques, hauts fonctionnaires, dirigeants d'entreprise, qui redoutent la trace difficilement réfutable de ce qu'ils auront dit.

Même lorsque quelqu'un a accepté, il n'est pas rare qu'il demande à l'enquêteur d'interrompre un moment l'enregistrement, par exemple lorsqu'il met en cause

nommément une autre personne, ou se livre à des critiques violentes. Le fait que, après la fin de l'entretien formel, et une fois le magnétophone débranché, une conversation, voire un monologue, apparemment plus libre puisse se poursuivre témoigne de l'effet inhibiteur qui peut intervenir. Mais il peut aussi être dû, en partie, à l'aspect très formel de la situation d'entretien, et pas uniquement à la présence du magnétophone.

Une expérience a permis de comparer des situations avec et sans magnétophone. Il s'agissait d'un questionnaire factuel, portant sur des points à propos desquels des recoupements permettaient d'apprécier l'exactitude des réponses. Globalement, les deux situations ont donné le même résultat, mais une analyse plus fine a montré que, en présence du magnétophone, les catégories sociales inférieures donnaient des réponses plus exactes que sans lui, alors que c'est l'inverse qui s'est produit pour les catégories supérieures.

Depuis peu de temps, on a essayé de compléter le magnétophone par le magnétoscope, qui permet d'enregistrer les gestes, la posture et la mimique. Mais jusqu'à présent on n'a guère trouvé de méthode permettant de tenir compte de ces informations complémentaires autrement que par des impressions qui restent toujours sujettes à caution.

5.1.4. Les questionnaires auto-administrés et les enquêtes par correspondance

Dans tout ce qui précède nous avons admis que c'est l'enquêteur qui pose les questions. C'est la méthode la plus souvent utilisée, celle qui permet le mieux de contrôler la situation. Mais on peut aussi remettre le questionnaire à l'enquêté, ou le lui adresser par la poste, en lui demandant de lire les questions et d'y répondre par écrit. En général, on choisit ce mode de passation pour des raisons d'économie : envoyer un questionnaire revient évidemment moins cher que de payer un enquêteur, mais cela présente des inconvénients importants, que nous verrons plus loin.

Les raisons financières ne sont pas les seules ; on recourt aussi au questionnaire auto-administré lorsque les questions peuvent embarrasser les sujets ; il vaut mieux alors leur éviter d'en parler, et les laisser répondre par écrit. C'est le parti qui a été adopté par l'I.F.O.P. pour la grande enquête sur le comportement sexuel des Français. Après une première partie constituée par des questions générales et des questions d'opinion, posées par l'enquêteur, celui-ci laissait à la personne enquêté un questionnaire portant plus précisément sur ses pratiques sexuelles. Une fois rempli, ce questionnaire était placé dans une enveloppe cachetée, et introduit dans une urne. La discrétion par rapport à l'enquêteur était ainsi sauvegardée, mais des repères permettaient, lors du dépouillement, de faire correspondre les deux questionnaires se rapportant à la même personne.

La même technique a été utilisée lors de sondages sur les intentions de vote, à la fois pour garantir l'anonymat des réponses, et pour mieux simuler la situation de vote. Mais, ici aussi, il faut pouvoir relier la réponse introduite dans l'urne à ce qu'on sait par ailleurs de la personne interrogée : son sexe, son âge, sa profession, sa position politique, etc. Sinon, on ne pourra obtenir que des résultats globaux, aucune analyse

n'étant possible. On le fait, sans que les sujets s'en aperçoivent, par la couleur des bulletins ou celle de l'enveloppe, ou par des signes discrets sur celle-ci. Mais il y a ainsi tromperie, même si on peut la juger sans gravité, et des sujets qui s'en sont aperçus ont suscité des incidents à ce propos.

Un questionnaire auto-administré ne peut pas être construit de la même manière que lorsque les questions sont posées par l'enquêteur. En effet, le sujet garde la possibilité de lire tout le questionnaire avant de commencer à répondre, de revenir sur une question à laquelle il a déjà répondu et de corriger sa première réaction, de laisser une question embarrassante en attente et d'y revenir plus tard, d'en parler éventuellement avec d'autres personnes, etc. Le problème du choix de l'ordre des questions se pose donc tout différemment. On ne peut plus, comme on pourrait le faire autrement, commencer par des questions générales, puis centrer l'attention sur des points particuliers, puisque les sujets peuvent très bien avoir vu les dernières questions avant de répondre aux premières. Certaines stratégies d'interrogation sont donc interdites par ce mode de passation.

Il faut aussi prendre garde au fait que, chez certaines personnes, le langage écrit n'est pas familier ; il ne s'agit pas seulement d'étrangers ne lisant pas bien la langue du pays, mais aussi de personnes de bas niveau d'instruction. Il faut donc un vocabulaire ou une syntaxe très simples. Pour la même raison, on évitera, dans la mesure du possible, des questions ouvertes, auxquelles on n'obtiendra que des réponses en majorité extrêmement sommaires. De même, on évitera les questionnaires en arbre, les questions filtres, tout ce qui peut donner lieu à erreur ou à malentendu.

Les instructions doivent être particulièrement précises. Par exemple, il faut indiquer très clairement si l'on demande de cocher la bonne réponse ou de biffer la mauvaise.

Pour faire une enquête par correspondance, ce qui est la forme la moins coûteuse, il faut disposer d'une liste d'adresses, donc d'une base de sondage. Si l'on est réduit à la méthode des quota, il faut au moins que des enquêteurs fassent du porte-à-porte ou accostent des personnes dans la rue, et leur laissent le questionnaire avec une enveloppe-réponse. Parfois, lorsqu'on procède à un échantillonnage sur place et que le questionnaire est très court, on peut recueillir les formulaires remplis quelques minutes plus tard.

Le principal inconvénient des enquêtes par correspondance, c'est le très faible taux de réponses. Celui-ci varie selon la population et selon l'intérêt de l'enquête, mais il peut descendre jusqu'à 10 %. Tous les biais que nous avons vus, à propos des refus ou des absences (cf. 2.2.4.2.), se retrouvent donc, aggravés. Des lettres de rappel peuvent être efficaces. Par exemple, dans une situation il est vrai particulièrement favorable (d'anciens étudiants interrogés par des chercheurs de leur Université au sujet de leur carrière), il a été possible d'obtenir près de 90 % de réponses avec 8 lettres de rappel. La méthode d'extrapolation des réponses recueillies par comparaison entre les personnes faciles à atteindre et les moins faciles, que nous avons vue à propos des absences, peut s'appliquer ici et permettre de redresser en partie les biais.

5.2. ESQUISSE D'UNE PSYCHOSOCIOLOGIE DE LA SITUATION D'ENQUÊTE

5.2.1. La situation d'enquête comme situation sociale

Une question est posée par un enquêteur à une personne désignée par le sort, ou choisie parce qu'elle présente une certaine caractéristique, et qui, par curiosité, par timidité, parce qu'elle n'avait rien d'autre à faire, pour avoir l'occasion de s'exprimer, parce que l'enquêteur lui était sympathique, ou simplement par inertie, a accepté de se prêter à l'enquête. Pour cela, elle a interrompu ses occupations habituelles, s'est éloignée de ses préoccupations du moment. L'enquêteur lui a posé des questions plus ou moins intéressantes, plus ou moins inattendues ou déroutantes. Arrivée à une certaine question, la personne interrogée a hésité, donné une première réponse, s'est reprise, a complété ce qu'elle avait dit en le corrigeant. L'enquêteur, lui, a mis une croix dans une case, éventuellement a noté quelques mots. Puis, tout le questionnaire a été réécrit sous la forme de chiffres inscrits dans un formulaire, transformé en trous dans une carte perforée, et introduit, avec les réponses de tous les autres sujets, dans un ordinateur d'où sortiront des tableaux de fréquences. Ce qu'a dit la personne interrogée est devenu un élément dans une série de distributions ou de tableaux croisés, avec beaucoup d'autres réponses. Elle est devenue une « donnée », qui est là, sous les yeux du chercheur. Comme le dit A. Touraine, on a fait comme si « les opinions étaient dans la tête comme les sous dans la poche » : il suffirait d'aller les y chercher [1].

Tout sociologue qui a eu l'occasion, ou le scrupule, de faire passer lui-même quelques questionnaires n'a pu manquer d'être frappé par le contraste entre le vague ou la complexité des réponses, et la netteté des distributions sur lesquelles il travaille ensuite, sur lesquelles on peut calculer des moyennes, des pourcentages, des corrélations, effectuer des analyses factorielles ou utiliser d'autres méthodes sophistiquées. On oublie ainsi que ces « données » ont été *produites* et non simplement *recueillies*, qu'elles sont l'aboutissement d'un processus en même temps que le point de départ, la matière première de l'analyse. Négligeant tout cet « avant », on dira donc couramment « 40 % de l'échantillon est allé au cinéma au moins une fois dans le mois qui précédait l'enquête » [2], et non « Au cours d'un questionnaire sur les sorties, à une question formulée de telle manière, 40 % des sujets ont répondu être allés au cinéma une fois ou plus, au cours du mois qui précédait l'enquête ».

Cet isolement de la réponse par rapport aux conditions dans lesquelles on l'a obtenue a des conséquences importantes. Dans la conception d'un questionnaire, ou lors de l'analyse des résultats, on raisonne en effet le plus souvent comme si tous les sujets interrogés avaient répondu « correctement », c'est-à-dire avaient toujours donné la réponse qui, à leurs yeux, était la plus juste. On admet facilement qu'un sujet puisse se

---

1. L. Churchill, analysant des enregistrements au magnétophone de passations de questionnaires fermés, a constaté de très importants écarts avec ce qui était « programmé », aussi bien de la part de l'enquêté que de l'enquêteur (20 % environ des réponses seulement étaient conformes à ce qui était prévu).

2. Cette formulation reste encore prudente. On trouvera souvent, surtout dans des présentations pour le grand public, « 40 % des Français vont au cinéma au moins une fois par mois ». Mais ce type de généralisation rapide n'est pas ce que nous visons ici.

tromper, que sa mémoire puisse être en défaut, éventuellement que la question ait pu être mal comprise ; on prend garde d'éviter, dans la conception du questionnaire, quelques erreurs grossières, mais la critique des réponses va rarement plus loin.

Pourtant, lorsqu'on prend la peine de vérifier, quand c'est possible, la validité de certaines réponses, on observe des « erreurs » curieuses. Par exemple, au cours d'une enquête sur les voyages d'affaire, on a confronté les réponses données par les sujets aux déplacements enregistrés par la comptabilité de leur entreprise. On a pu ainsi constater qu'à côté d'omissions ou d'erreurs sur les fréquences, qu'il est peut-être possible d'attribuer à des défaillances de la mémoire, certains voyages ont été mentionnés, qui, d'après tout les recoupements, n'ont jamais été effectués, leur nombre atteignant près de 20 % du nombre total de voyages cités. D'autres travaux, sur lesquels nous reviendrons plus loin, confirment l'existence, et l'importance, de distorsions qu'il est difficile de considérer comme de simples erreurs aléatoires sans signification.

A la suite de telles observations, la confiance dans la validité des réponses obtenues par questionnaire a été sérieusement ébranlée. Si des distorsions se produisent sur des points aussi clairement définis qu'un voyage, qu'en sera-t-il des questions portant sur des opinions, beaucoup moins bien définies ? On a commencé à prendre conscience du fait que le déroulement d'un entretien, ou la passation d'un questionnaire constituent des situations sociales, dans lesquelles entrent en interaction au moins deux personnes, l'enquêteur et l'enquêté, auxquelles peut éventuellement s'ajouter la présence virtuelle, mais bien agissante, de l'organisation enquêtrice ou de son client, sous la forme des représentations, plus ou moins correctes, que l'enquêté s'en fait.

Entre la sérénité naïve de certains bureaux d'études qui ignorent totalement le problème, et le rejet de ceux qui affirment péremptoirement qu'un enquêteur habile peut faire dire n'importe quoi à n'importe qui, il y a place pour des recherches psychosociologiques empiriques sur la situation d'enquête, considérée comme une situation sociale, qui n'est pas une donnée, qui n'est pas nécessairement ce que le chercheur a décidé qu'elle devait être, mais qui se construit, se définit, se « négocie », au cours de son déroulement. Depuis quelques années, un certain nombre de ces recherches tentent de débroussailler le problème, de mettre en évidence les biais éventuels qui peuvent s'introduire, d'en évaluer l'importance, et de donner des moyens de les éviter. Même si les conclusions de ces recherches ne sont pas encore tout à fait claires, il s'en dégage nettement que la situation d'enquête n'est pas une situation « neutre » (d'ailleurs, est-ce que cela aurait vraiment un sens ?), et que les réponses ainsi recueillies doivent être, par conséquent, toujours interprétées en référence à cette situation.

La psychologie expérimentale s'est posé le problème depuis plus longtemps, et la situation d'expérience en laboratoire a été étudiée de façon approfondie. On a ainsi montré que la situation définie par l'expérimentateur n'est pas nécessairement celle qui est perçue et vécue par le sujet. Par exemple, du point de vue du chercheur, un expérimentateur a exposé son sujet à une source d'influence, et attend de lui qu'il prête attention au message émis, puis indique sérieusement son opinion. Pour le sujet, ce peut être tout à fait autre chose : il participe à une expérience montée par son professeur de psychologie sociale, il fait des hypothèses sur l'objet de cette expérience et ce qui en est attendu, il cherche peut-être à briller, à en tirer des enseignements, à se montrer dans la

norme, à prouver que la méthode expérimentale ne vaut rien, il redoute un piège et cherche à se montrer le plus malin, etc. Là où l'expérimentateur lui fixe comme objectif d'être sincère dans ses réponses, le sujet, s'il est de bonne volonté, pourra chercher à être le plus coopératif possible, ce qui n'est pas nécessairement la même chose. Les deux situations, celle qui a été conçue par l'expérimentateur, et celle qui est perçue par le sujet, ont peu de chance de coïncider. Elles pourront être suffisamment proches, parfois, pour ne pas invalider entièrement les conclusions, mais ce n'est pas certain. De plus, on a montré que l'expérimentateur, malgré toute sa volonté de neutralité, peut communiquer involontairement au sujet ses hypothèses et ses attentes, renforçant sans s'en rendre compte, par son attitude, le ton de sa voix, ou d'autres indices encore plus fugaces, certains comportements au détriment d'autres.

Ces conclusions, dont nous n'avons esquissé que les très grandes lignes, amènent à poser le problème dans les termes généraux suivants : *dans une expérience, ou lors de la passation d'un questionnaire, quelle est la situation pour le sujet, quels sont les objectifs qu'il poursuit dans cette situation, et quelles sont les stratégies qu'il adopte pour atteindre ces objectifs ?*

L'idéal, pour le praticien des enquêtes, n'est pas de pouvoir susciter une situation « neutre » où les réponses ne seraient pas biaisées par le contexte, mais une situation telle que la meilleure stratégie pour le sujet, ou la plus probable, consiste à donner la réponse qui lui semble la plus exacte. Si cet idéal ne peut pas être atteint avec certitude, la compréhension de la situation pourra permettre de mieux interpréter les réponses, d'éviter de les prendre trop au pied de la lettre, et finalement de faire des hypothèses quant aux biais qui auraient pu les affecter.

Dans les enquêtes, la définition de la situation est vraisemblablement plus simple que dans le cas des expériences de laboratoire, qui sont souvent artificielles, et au cours desquelles on recourt à des compères, etc. Pour la majorité de la population, la procédure qui consiste à aller demander aux gens ce qu'ils pensent est assez claire et normale, et on parle suffisamment des sondages d'opinion pour que le fait de voir un enquêteur sonner à la porte ou aborder quelqu'un dans la rue, ne paraisse pas incongru ou incompréhensible, même si l'on est, pour des raisons diverses, opposé aux sondages et si on les juge indiscrets. Cela n'empêche pas toutefois certains malentendus : pour celui qui a conçu l'enquête, l'objectif est souvent à la fois large et complexe, surtout s'il s'agit d'une recherche, et non d'une étude visant à résoudre un problème bien circonscrit ; l'enquêté, lui, imagine en général des buts plus précis, plus immédiatement utiles, plus proches de la manière dont lui-même pose le problème traité. A la limite, on peut dire que ce sont deux conceptions de la sociologie, en tant que science et en tant que pratique, celle du chercheur et celle, implicite et informelle, du profane, qui sont confrontées. Du point de vue du sujet, certaines questions paraîtront complètement inutiles, oiseuses, et le dérouteront, parce qu'il ne comprendra pas pourquoi on les lui pose. S'il a une conception très utilitaire de l'enquête, il pourra saisir l'occasion d'exprimer ses doléances, ou ses revendications. Par-delà l'enquêteur, il s'adressera aux autorités ou à ceux qui, à ses yeux, ont pouvoir sur sa situation. Dans ces conditions, sa stratégie ne l'amènera pas nécessairement à donner la réponse « correcte », mais, par exemple, à noircir la description qu'il fera de sa situation.

5.2.2. Les rôles de l'enquêté

Un des aspects de la situation telle qu'elle est vécue par le sujet, c'est le *rôle* qu'il adopte ou croit devoir adopter. On peut répondre à un même questionnaire en tant que salarié, ou en tant que père de famille, ou en tant qu'électeur, etc. Selon le cas, les réponses qu'on donnera pourront ne pas être les mêmes. Ces différents rôles ne sont évidemment pas propres à la situation d'enquête ; chacun joue plusieurs rôles, et agit chaque fois en accord avec celui-ci. Le problème, lors de la conception de l'enquête, n'est donc pas de chercher à éliminer l'effet de rôle, mais de savoir lequel des rôles est le plus pertinent pour le problème traité.

Par exemple, au cours d'une étude sur les besoins en équipements pour enfants, comprenant une enquête auprès des mères de famille d'une commune de la banlieue parisienne, presque toutes les personnes interrogées ont déploré vigoureusement l'absence de tels équipements dans leur quartier et l'impossibilité où elles étaient de quitter leurs enfants pour une heure ou deux. Or, à peu près au même moment, une assemblée de co-propriétaires, à laquelle avaient participé la plupart des femmes interrogées, avait examiné, et rejeté à la quasi-unanimité, une proposition de créer, dans un local inoccupé, une garderie pour les enfants. La contradiction entre les réponses à l'enquête et le comportement au cours de l'assemblée peut s'expliquer, entre autres, par une différence de rôles. Les enquêteurs abordaient les habitantes de l'ensemble en leur demandant si elles avaient des enfants, puisque c'était là la définition de la population étudiée. Elles savaient donc que c'était en tant que mères de famille qu'elles étaient interrogées, et répondaient en conséquence. Lors de la réunion, par contre, ces mêmes personnes avaient réagi en tant que co-propriétaires, sensibles aux problèmes de coût, aux difficultés de gestion, de relations entre voisins, etc. Le rôle dans lequel elles avaient été interrogées n'était pas celui dans lequel elles devaient prendre la décision. Cela ne veut pas dire que leurs réponses doivent être considérées comme étant sans valeur : les conclusions qu'on en a tirées concernant les problèmes de garde d'enfants sont probablement correctes. Mais ce faisant on ne recouvrait pas la totalité des aspects du problème, et en particulier on aurait eu tort d'extrapoler, et de prévoir des comportements.

Le rôle dans lequel se place une personne interrogée est probablement déterminé, en partie, par la manière dont elle est abordée par l'enquêteur. Très souvent, en particulier lorsqu'on vise une sous-population particulière, ou lorsqu'on effectue un échantillonnage par quota, on est obligé de poser, d'emblée, les questions nécessaires pour savoir si la personne doit être interrogée ou non. On lui communique ainsi que c'est en tant que présentant ces caractéristiques qu'elle est choisie et qu'on lui demande son avis. Il y a des chances qu'elle réponde en conséquence.

Une manière d'éviter ce biais, si on estime que c'en est un, c'est de poser plusieurs questions, même inutiles pour le choix des sujets, afin que ceux-ci ne puissent pas savoir quelle est la caractéristique qui est importante, et donc choisir eux-mêmes le rôle en fonction duquel ils répondront. Ou alors, pour éviter cette indétermination et mieux maîtriser la situation, on renforce l'indication d'un certain rôle, de manière que cet aspect soit commun à tous, et qu'on puisse le faire intervenir dans les interprétations.

5.2.3. Les objectifs de l'enquêté

Les raisons qui amènent quelqu'un à accepter de se prêter à un entretien, ou à répondre à un questionnaire, peuvent être très diverses, allant de la timidité qui empêche de refuser, à l'intérêt scientifique, en passant par le désir d'avoir un contact un peu inhabituel, le désir de donner son avis, ou la fierté de participer à un sondage dont il sera peut-être question dans le journal. Le comportement durant la passation est lui aussi déterminé par des objectifs divers. Les recherches effectuées au cours des dernières années ont permis de mettre en évidence deux catégories de ces objectifs, les uns liés à la relation avec l'enquêteur, les autres liés à l'utilisation qui, dans la représentation du sujet, pourra être faite de ses réponses. Il est peu probable que ces deux catégories épuisent le domaine, mais on touche certainement ici des points particulièrement importants.

En ce qui concerne les *relations avec l'enquêteur,* sur lesquelles on reviendra plus loin, on peut distinguer trois objectifs, d'ailleurs étroitement liés :

— *Maintenir de bons rapports avec l'enquêteur,* c'est-à-dire ne pas le heurter, donner des réponses qu'il pourrait désapprouver et risqueraient de susciter chez lui une réaction désagréable. On peut probablement expliquer par l'existence de cet objectif la constatation qu'il existe, dans certaines conditions, une corrélation, faible mais néanmoins significative, entre les opinions d'un enquêteur et celles qu'il obtient lorsqu'il fait passer un questionnaire. Au cours d'une expérience systématique, on a également observé que, lors de la passation d'un questionnaire d'attitude politique, un simple *right* approbateur de l'enquêteur à chaque réponse d'un certain type entraînait un accroissement significatif (de l'ordre de 25 % en moyenne) du nombre de réponses de ce type.

— *Donner de soi une image favorable.* Ce peut être un des moyens d'obtenir l'approbation de l'enquêteur, mais aussi, peut-être, une fin en soi, puisqu'on rencontre également des biais dans ce sens dans les questionnaires auto-administrés. L'action de cet objectif a été vérifiée au moyen de questionnaires portant sur des faits qu'il était possible de connaître par ailleurs, tels que d'avoir voté lors d'une certaine élection, ou les résultats scolaires des enfants. On a observé une déformation systématique pouvant entraîner une surestimation allant jusqu'à 30 % des comportements les plus valorisés.

— *Donner de soi une image « conforme », normale.* Cet objectif est très difficile à distinguer du précédent, mais il semble que le souci d'apparaître « dans la norme » soit, dans certains cas, plus fort que le désir de se valoriser. On peut l'interpréter soit comme un moyen de défense contre une utilisation abusive ou indiscrète des résultats, soit comme une stratégie pour maintenir des rapports satisfaisants avec un enquêteur dont on ignore à peu près tout.

L'autre source d'objectifs des enquêtés, eux aussi susceptibles de biaiser les réponses, réside dans l'adoption de réponses « *instrumentales* », expressions d'une *stratégie* pour atteindre un but extérieur à l'enquête qui devient un moyen pour le sujet. Une personne interrogée sur ses conditions de logement par exemple, et qui estime que l'Etat n'en fait pas suffisamment dans ce domaine, pourra noircir sa situation, espérant ainsi inciter

les autorités à prendre les mesures qu'elle souhaite. On voit que ce qui est important, c'est la représentation que la personne interrogée se fait de l'utilisation future de ses réponses.

L'existence simultanée de ces différents objectifs a été montrée dans une enquête très ingénieuse menée dans une ville américaine où un service public était en grève. Lors d'un sondage d'opinion sur la grève, une moitié des enquêteurs se sont présentés comme faisant partie d'un institut de sondages, donc « neutres », les autres comme faisant partie du syndicat qui menait la grève. Le questionnaire comportait à la fois des questions portant sur la catégorie de fonctionnaires concernés, visant à estimer leur popularité, et d'autres portant plus spécifiquement sur la grève elle-même.

Les résultats sont, à première vue, paradoxaux : les enquêteurs « syndicaux » ont obtenu à la fois des réponses indiquant une popularité plus grande que les autres enquêteurs, et une hostilité plus marquée à la grève. Vraisemblablement, aux questions demandant des jugements généraux sur les fonctionnaires, c'est le souci d'avoir un bon contact avec l'enquêteur qui a prédominé ; en revanche, lorsqu'on est passé aux questions sur la grève elle-même, certaines personnes ont saisi l'occasion de faire savoir au syndicat qu'elle les gênait et qu'elles souhaitaient qu'elle cesse.

5.2.4. L'influence de l'enquêteur

Tout ce que nous venons de dire montre l'importance de la relation enquêteur-enquêté. La manière dont les objectifs des enquêtés influencent leurs réponses ne peut se comprendre que par rapport à cette relation, en se souvenant que l'enquêteur, dans cette situation, ne se réduit pas à sa personne propre, mais qu'il est aussi le délégué d'une institution vis-à-vis de laquelle peuvent s'exercer diverses stratégies. Inversement, l'influence de l'enquêteur ne peut se comprendre que par référence à ces objectifs.

Deux facteurs ont été particulièrement étudiés à ce propos : la distance sociale entre l'enquêteur et l'enquêté, et la qualité du contact. A priori, on pourrait penser, et c'est ce qui a été longtemps admis, qu'il était préférable que cette distance soit minimum, donc que l'enquêteur soit aussi semblable que possible à ceux qu'il doit interroger, et que le contact soit aussi bon que possible.

Un ensemble de recherches expérimentales récentes, au cours desquelles on a fait varier systématiquement ces variables, a donné des résultats confus dans leur détail, mais où deux points ressortent nettement : les réponses les plus exactes (lorsqu'il est possible de contrôler leur véracité) ne sont obtenues ni lorsque la distance sociale est minimum, ni lorsque le contact entre l'enquêteur et l'enquêté est le meilleur. Une certaine distance, à la fois sociale et affective, est nécessaire. Cela peut probablement s'expliquer par le fait que le désir de maintenir un bon contact avec l'enquêteur, de lui donner une bonne image de soi, est probablement maximum lorsque cet enquêteur est proche. Lorsque la distance augmente, au moins jusqu'à un certain point, on deviendrait plus indifférent à ses réactions.

Ces conclusions doivent être nuancées : l'interaction entre la distance sociale et la qualité du contact donne des effets complexes, qui, en outre, varient selon les catégories sociales. De plus, contrairement à ce qui a été constaté en ce qui concerne l'influen-

ce des objectifs des enquêtés, qui était massive, l'effet différentiel des caractéristiques des enquêteurs est beaucoup plus faible.

Malgré toutes ces réserves, les points de convergence des différentes recherches invalident ce qu'on pourrait appeler le *modèle de la confidence,* sur lequel on pourrait être tenté de régler les conditions d'interview. *Un entretien ou la passation d'un questionnaire ne constituent pas une relation amicale ni une relation d'aide,* même si certains sujets le souhaitent, et même si ce serait plus agréable pour les enquêteurs. C'est une technique de recueil d'informations. Si, effectivement, on peut penser qu'il y a des problèmes qu'on n'aborde qu'avec des proches, et pas avec un enquêteur, rien ne nous assure que ce qui est dit au cours de ces conversations correspond à la « vérité » que voudrait obtenir le chercheur.

Cela amène à examiner d'un œil critique certaines attaques formulées contre les enquêtes, et qu'on peut résumer par : Que peut-on espérer obtenir par l'intermédiaire d'un inconnu qui applique des consignes stéréotypées ? De ce point de vue, seule l'observation participante, ou l'auto-observation des membres d'un groupe ou d'une institution, permettrait d'obtenir des informations et des réactions qu'on refuse à un étranger rencontré pour un temps très bref. Mais on oublie alors que, dans bien des cas, c'est précisément parce que l'enquêteur n'est qu'un étranger de passage que les stratégies à son égard pourront être minimales.

Ceci dit, il ne faut pas croire non plus qu'on puisse tout obtenir par enquête. Bien que l'idée que la distance sociale *optimale* entre enquêteur et enquêté n'est pas la distance *minimale* soit fortement suggérée par les recherches que nous avons mentionnées, elle rencontre certainement des limites.

En premier lieu, les normes de la plupart des groupes sociaux imposent qu'on ne parle pas de certains thèmes avec des personnes « neutres » ou mal connues ; cela implique qu'on ne peut en général recueillir des informations à leur propos que par l'intermédiaire d'une relation plus personnalisée, que l'on maîtrise très mal, et qui a de fortes chances de biaiser les informations fournies. Dans de tels cas, il faut soit utiliser des techniques détournées, comme le questionnaire auto-administré de l'enquête sur le comportement sexuel des Français (dont nous ignorons la validité), soit accepter les risques de biais dus à des relations trop personnelles, et en tenir compte dans l'interprétation.

BIBLIOGRAPHIE

La psychosociologie des relations entre chercheur et sujet, dont la situation d'enquête est un cas particulier, a été surtout étudiée dans le cas des expérimentations de psychologie en laboratoire. Quelques textes importants sont présentés dans :

- G. LEMAINE, J.M. LEMAINE, *Psychologie sociale et expérimentation,* Paris-La Haye, Mouton, 1969.

Pour une synthèse plus complète :

- R. ROSENTHAL, R.L. ROSNOW, *Artifact in Behavioral Research,* New York, Academic Press, 1969.

En ce qui concerne plus spécifiquement les enquêtes :

- D.L. PHILIPS, *Knowledge from what ?* Chicago, Rand MC. Nally, 1971.

ainsi que de nombreux articles et discussions dans la revue *Public Opinion Quarterly,* principalement depuis 1968.

Deuxième partie

L'ANALYSE ET L'INTERPRÉTATION DES DONNÉES D'ENQUÊTES

Chapitre 6

L'ANALYSE DE CONTENU

6.1. GÉNÉRALITÉS

Un grand nombre d'ouvrages ont été publiés, dans la dernière décennie, sur le problème des carences de l'analyse de contenu, tant au plan méthodologique qu'au plan théorique. Une analyse de contenu sommaire et thématique (ou thématique donc sommaire ?) de quelques-unes des définitions proposées montre une dégradation de l'image « de marque » de cette technique :

« L'analyse de contenu est la statistique sémantique du discours politique » (Kaplan, 1943).

« L'analyse de contenu est une technique de recherche pour la description objective, systématique, et quantitative du contenu manifeste de la communication » (Berelson, 1952).

« On propose d'utiliser le terme d'analyse de contenu et de codage de façon interchangeable pour renvoyer à la description objective, systématique, quantitative de toute conduite symbolique » (Cartwright, 1953).

« L'analyse de contenu réfère à toute procédure pour fixer l'extension relative qui spécifie référents, attitudes ou thèmes contenus dans un message ou un document » (Stone, 1964).

« L'analyse de contenu est une technique pour faire des inférences par identification systématique et objective des caractéristiques spécifiques d'un message »,

« L'analyse de contenu est un ensemble disparate de techniques utilisées pour traiter des matériaux linguistiques »,

« L'analyse de contenu est un instrument analytique sans ambition sinon sans fondements théoriques » (P. Henry, S. Moscovici, 1968).

« L'analyse de contenu résout empiriquement le problème de l'articulation de la langue à la situation, mais... elle résout ce problème par l'intermédiaire d'une situation théorique réellement intenable en tant qu'elle ne possède pas le concept de sa pratique » (Trognon, 1972).

Enfin, une critique radicale de l'analyse de contenu pourra être trouvée dans le numéro 37 de *Langages* sous la plume de Pecheux (1975).

Mais il n'en reste pas moins que la pratique de l'analyse proposée par Pecheux est toujours plus ou moins entachée des défauts constitutifs qui ont été ceux de toute analyse de contenu : l'intervention du codeur dans l'établissement du sens du texte. Cette distance entre pratique et théorie peut être présentée comme un « écart » à combler entre une « analyse du discours » qui serait imparfaite et une « théorie du discours » qui le serait moins, mais cela peut également être relié à une pratique qui n'a pas sa théorie ou qui n'a que des réponses partielles, dont rien ne peut garantir qu'elles seront les bases d'une théorie générale, ce qui serait beaucoup plus grave.

Ainsi, à sept ans de distance (1968 : Moscovici ; 1975 : Pecheux) on est toujours confronté au même problème, bien qu'un certain nombre de déplacements aient eu lieu (de l'analyse du contenu à l'analyse du discours « l'AAD visant elle-même à détruire de ce point de vue l'analyse de contenu » Pecheux, 1975) et que les théories et les techniques aient progressé, celui d'une pratique non fondée point par point par une théorie. Spécifique reste l'interrogation de la possibilité d'une telle adéquation ? A cette question les réponses sont diverses, d'aucuns croient à la « révolution » qui apportera une solution : une « transformation radicale du problème dans ses termes mêmes », d'autres croient à la continuité que représenterait une sémantique générative, certains enfin sont perplexes.

De ce qui nous apparaît aujourd'hui — sans qu'il soit préjugé en aucune façon de l'avenir — une difficulté majeure, on dégagera dans le cadre de cet ouvrage deux conclusions pratiques :

— Le but que nous poursuivons est de permettre au *praticien* de se servir de ce qui est dit ; par conséquent, même si les problèmes théoriques ne sont pas ignorés, ils ne peuvent en aucune façon condamner au silence. Car si le paradoxe veut que l'absence théorique permette à beaucoup plus de gens de parler pour la déplorer que ne le ferait une réponse théorique satisfaisante, le but de cet ouvrage n'est pas d'ajouter aux lamentations un paragraphe supplémentaire, même si ces dernières font évoluer la science.

— La pratique de l'analyse de contenu, considérée comme une méthode *et cela seulement,* est avant tout une pratique sociale. Elle trouvera ses justifications dans une prédictibilité empirique liée à un objectif prédéterminé et cherchera des validations extrinsèques à sa propre pratique.

On sait combien une telle affirmation prête le flanc aux critiques théoriques et on voudrait signifier, ce disant, qu'on ne l'ignore pas. On les partage pour un grand nombre d'entre elles mais, ainsi que le soulignent plusieurs détracteurs, « l'analyse de contenu fonctionne comme pivot dans l'ensemble des sciences sociales ».

Afin de préciser ce point qui nous paraît crucial — et avant d'entamer le discours méthodologique proprement dit — on voudrait rapidement présenter une synthèse des demandes sociales auxquelles ont eu à répondre différentes sortes d'analyses de contenu (et non pas seulement les analyses de contenu d'entretiens).

Holsti (1968) soulève cette question mais ne lui donne pas cet aspect insistant de demande sociale qui semble être une des caractéristiques fondamentales de l'analyse de contenu, et qui est à l'origine de sa faiblesse théorique. Répondant à une demande sociale celle-ci ne pouvait guère développer qu'une méthode, la plus fiable possible, pour satisfaire des exigences de coût et d'objectivité. D'où l'accent mis sur l'objectivité, la systématicité, la capacité de généralisation. Une telle attitude ne pouvait, à son tour, que produire une fuite en avant méthodologique, c'est-à-dire un raffinement toujours plus grand de l'outil, afin de le rendre plus fiable, toujours mieux adapté à la demande.

Une des demandes les plus connues [1] est celle émanant de la *Federal Communications Commission* de la *Library of Congress* et du *Justice Department* qui, durant la deuxième guerre mondiale, engagèrent un grand nombre de chercheurs pour étudier la propagande ennemie et amie. On comprendra aisément à quelle rationalité correspondait une telle demande. Il était en effet important pour les Etats-Unis de connaître les méthodes et les thèmes de la propagande ennemie afin de bâtir une contre-propagande ou d'abolir les sources de propagande. Il était de même important de connaître les mécanismes de l'influence sociale exercée par un texte (ou un film, ou un discours, ou une bande dessinée) afin de pouvoir exercer, par des actions de propagande interne, une pression sur les Américains, dans le sens souhaité par les pouvoirs publics, pour soutenir l'effort de guerre.

Un autre type de demande, de nature clinique, recherche les traits psychologiques des individus à travers leurs écrits, récits, etc. Là encore la demande est aisément compréhensible quant à sa rationalité. Dans une optique thérapeutique, l'analyse de contenu aidera à fournir un meilleur diagnostic et, éventuellement, pourra servir d'outil thérapeutique. Cette demande, de loin la plus importante, part du postulat selon lequel des troubles psychiques trouvent un de leurs équivalents comportementaux dans les troubles du processus de communication. Dans une autre optique, elle pourra servir à prédire le comportement d'individus dans telle ou telle circonstance et par rapport à tel objet, ou idée, ou situation.

Un troisième type de demande liée à l'apparition conjuguée du travail en groupe et de la psychothérapie de groupe, elles-mêmes liées l'une à l'échec du travail individuel pour résoudre les problèmes complexes (ou à la plus grande rentabilité du travail en groupe, ce qui a pour conséquence de relativiser la notion d'échec), l'autre à l'apparition d'un courant issu des théories moréniennes, conduit à mettre au point des métho-

1. On citera ici surtout des sources américaines, les sources françaises étant relativement rares.

des d'analyse des interactions de groupe. La méthode de Bales (1950) est l'une des plus connues [1].

Parallèlement, et de façon « logique », une autre demande se faisait jour, celle de l'analyse de la prise de décision dans des circonstances données. Dans le même ordre d'idées, les demandes économiques (études de motivation) ou industrielles (analyse des entretiens du personnel en vue de modifier une structure) voyaient leur nombre augmenter.

Il semble donc, au moins a priori et pour l'essentiel, que les trois grands ordres de demandes soient politiques, économiques et thérapeutiques.

On essaiera de regrouper dans un tableau (p. 159) quelques-unes des différentes demandes spécifiques auxquelles ces trois ordres ont donné lieu quant à l'analyse de contenu.

Ce rapide survol des domaines d'application de l'analyse de contenu était destiné à en marquer l'étendue et, par conséquent, à justifier la place centrale que ces méthodes occupent dans les sciences sociales.

Toutefois, notre propos n'est pas d'analyser l'ensemble des méthodes possibles, mais plutôt de nous intéresser aux analyses de contenu effectuées sur le matériel d'enquête habituel : les entretiens.

Dans ce chapitre on tentera de rester au plus près du déroulement effectif d'une enquête et, par conséquent, de répondre successivement aux questions qu'on peut se poser lorsqu'on désire introduire dans les modalités de traitement une ou plusieurs méthodes d'analyse de contenu.

La structure de ce chapitre sera donc liée à une chronologie des interrogations du chercheur, qui pourrait être la suivante :

— Est-ce que le problème que je me pose est passible d'un traitement relevant de l'analyse de contenu ?

— Quelles sont les méthodes d'analyse de contenu ?

— Comment s'appliquent-elles ?

— Comment tester leur fidélité, leur validité ?

— Quels sont les types d'analyse de contenu dont on dispose dans le cadre d'une enquête ?

Cet ordre, purement empirique, n'empêchera nullement de poser les problèmes théoriques ou méthodologiques que chacune des étapes suscite. Nous éliminerons volontairement l'aspect historique et nous ferons un choix dans les méthodes, car prétendre à l'exhaustivité en ce domaine relève du pur phantasme.

6.2. PROCÉDURES CLOSES ET OUVERTES

Une catégorisation commode est celle de Henry et Moscovici (1968) qui distinguent des procédures closes et des procédures exploratoires ou ouvertes.

La première « consiste à partir d'un cadre empirique ou théorique d'analyse de cer-

1. On ne développera pas par la suite les différentes méthodes d'analyse des relations de groupe, relativement extrinsèques à l'objet propre de ce chapitre.

TABLEAU I

DOMAINES D'INTERVENTION DE L'ANALYSE DE CONTENU

1. Politique Sociaux	Analyse de la propagande ; étude des processus d'influence sociale ; études globales de sociétés en fonction du « *need of achievement* », études globales de sociétés en fonction du critère « *inner/other directions* » ; analyse de textes destinée à fournir une expertise légale (par exemple, pour déterminer si tel journal diffuse de la propagande ennemie) ; études portant sur la prise de décisions en politique étrangère.
2. Economique	Etudes de motivation ; études structurelles au sein d'entreprises ; analyse du fonctionnement de groupes devant résoudre un problème ; analyse de la prise de décisions (en situation normale, en situation d'urgence, etc.) ; analyse de la circulation d'informations au sein d'une structure ; analyse de tâches ; analyse de comportements d'achats ; analyse de perception de publicité, etc.
3. Thérapeutique	Analyse de la personnalité, « *personnal structure analysis* », analyse des idéologies ; ensuite un certain nombre d'analyses spécifiques donnant lieu à mesure : DRQ *(Discomfort Relief Quotient)* destiné à tester le nombre de mots « *discomfort* » par rapport aux mots « *comfort* » dans un discours et à suivre leur évolution ; le PNAvQ *(Positive Negative Ambivalent Quotient)* proche du DRQ ; le TTR *(Type Token Ratio)* qui mesure la variété du vocabulaire ; l'AV *(Adjective Verb)* qui mesure le nombre d'adjectifs pour n verbes (liés à la distinction schizophrènes - normaux), mesure des désordres du discours, de l'anxiété, de l'hostilité, etc.

tains états psychologiques, psychosociologiques, ou autres, qu'on cherche à particulariser ou bien à propos desquels on a formulé des hypothèses ou des questions. On rassemble des textes produits dans des conditions telles que ces états fassent partie de leurs conditions de production. Puis *on observe* [1] ces textes à travers le cadre théorique fixé pour aboutir à une particularisation des conditions de production considérées ou à une

1. C'est nous qui soulignons.

mise à l'épreuve des hypothèses les concernant. Ces procédures sont dites closes car le cadre pré-établi ne peut être modifié ».

La seconde consiste à « rassembler des textes exploratoires produits dans des situations particulières correspondant aux états qu'on veut observer ; le cadre d'analyse de ces dernières n'est pas fixé et on part de la mise en évidence des propriétés des textes. On définit des différences, des similitudes et éventuellement des transformations faisant se correspondre les textes ou leurs représentations. Les différences, les similitudes, les transformations, doivent ensuite être interprétées pour aboutir à une caractérisation des états observés ».

Cette dichotomie marque, semble-t-il, nettement la différence entre les deux grands types d'analyse de contenu taxinomiques [1] et peut parfaitement être considérée comme cohérente avec la définition de l'analyse de contenu fournie par Holsti (1968) : « L'analyse de contenu est une technique pour faire des inférences, par identification systématique et objective des caractéristiques spécifiques d'un message. »

Sans doute ces procédures constituent « des techniques d'observation d'états qu'on ne cherche pas à appréhender en tant que conditions de productions de textes... mais en quelque sorte en eux-mêmes, indépendamment du fait qu'ils peuvent intervenir dans le processus de production » (Henry et Moscovici, 1968). Sans doute est-ce là une limite importante à « leur rentabilité théorique ». Mais en l'absence d'une réponse satisfaisante tant au plan théorique qu'expérimental ou pratique, des méthodes d'analyse de contenu systématiques (c'est-à-dire celles qui considèrent que les phénomènes étudiés et les processus de production ont leurs lois propres, que ces deux ordres de phénomènes sont en relation dans le jeu du croisement du plan horizontal au plan vertical et réciproquement et que le codage — phénomène essentiel de l'analyse de contenu — est soluble par une théorisation de cette relation), il semble difficile d'attacher une importance définitive à ces critiques, en tout cas, une importance suffisante pour annuler dans l'immédiat les méthodes d'analyse de contenu taxinomiques et permettre leur remplacement par des méthodes systématiques.

6.3. CHAMP D'APPLICATION DE L'ANALYSE DE CONTENU DANS UNE SITUATION D'ENQUÊTE

Par souci de clarté, on partira du schéma interrogatif classique « Qui dit, quoi, à qui, comment, pour quel effet ? ». Chacune des parties de la question, composant elle-même une question indépendante des autres, peut renvoyer à un ou plusieurs des trois buts possibles suivants :

— Analyser les caractéristiques du texte en tant que tel (« plan horizontal ») [2].
— Analyser les causes et antécédents du message (« plan vertical »).
— Analyser les effets de la communication.

---

1. Il faut entendre par-là, selon Henry et Moscovici (1968), « des indices des états permettant de les classer (ces textes) en établissant entre eux des différences ou des similitudes ».

2. Qu'on veuille bien excuser ce que l'on sait être une approximation conceptuelle.

La première démarche donne lieu à trois comparaisons :

— celle de documents issus d'une même source,
— celle de documents issus de différentes sources,
— celle de documents avec des catégories externes, par exemple des concepts psychanalytiques.

La seconde renvoie à Osgood (1959) selon lequel, dans ce cas, on cherchera « les lois des relations entre les événements dans les messages et les processus sociaux médiatisés par l'individu qui les produit », ce qui revient à chercher les conditions de production des textes et à éviter de mettre l'accent sur le seul individu producteur (considéré comme une entité indépendante de toute procédure sociale).

Deux processus de comparaison peuvent alors être envisagés :

— Direct, les données seront issues de l'analyse de contenu et comparées avec des indices comportementaux pris dans d'autres domaines que le linguistique.

— Indirect, les données seront également comparées à des indices comportementaux pris dans d'autres domaines, mais par le biais d'un raisonnement syllogistique : si, dans une situation x, un individu produit des messages présentant les caractéristiques x1, x2, x3 et x'1, x'2, x'3, considérées comme deux familles de caractéristiques différentes, correspondant à deux comportements (1 et 2),

: si, dans une situation x' semblable à la situation x, un individu produit des messages présentant les caractéristiques x'1, x'2, x'3,

: alors le message produit correspondra au comportement 2.

Le troisième renvoie aux analyses liées au champ de l'influence sociale et peut être formulé de la façon suivante :

— soit un individu A dont les caractéristiques comportementales le conduisent à produire un message x par rapport à un objet y,

— soit un individu B dont les caractéristiques comportementales le conduisent à produire un message x' destiné à A, par rapport au même objet y,

si, après avoir entendu le message de B, A produit un message présentant les caractéristiques du message x', alors le message pourra être référé au comportement de B.

Toutefois, il faut bien être conscient qu'à première vue, *dans le cas de l'analyse d'entretiens,* la question la plus importante apparaît être le « quoi ? ». Etant donné cette situation particulière, le « quoi ? » fournit en principe d'emblée la possibilité de répondre à toutes les autres questions. On sait « qui » a parlé et « à qui ». Le « comment ? » et « par quel effet ? » sont moins bien définis, mais ce devrait être relativement clair : le sujet parle parce qu'on le lui a demandé, et la réponse au « comment ? » se trouve dans le discours lui-même. Néanmoins, nous l'avons vu plus en détail dans le chapitre consacré à la psychosociologie de la situation d'enquête, ces évidences demandent à être réexaminées.

Le sujet peut parler du point de vue de l'un des différents rôles qu'il lui arrive de jouer ; sa représentation de l'enquêteur n'est pas une donnée, mais se construit au cours de leur interaction. Quant à ses objectifs (pour quel effet ?) ils peuvent être très divers. Toutefois, ces questions débordent largement la seule analyse de contenu des entretiens bien que celle-ci puisse largement aider à leur trouver une réponse. Il reste

donc, spécifiquement, la question « quoi ? » à laquelle d'ailleurs toutes les autres peuvent être ramenées, puisque, dans les cas où l'analyse de contenu s'estime en mesure d'y répondre, c'est toujours par l'intermédiaire du « quoi ? ».

Pourtant, poser, à propos d'un texte ou d'un discours, et pas seulement d'ailleurs à propos d'un entretien, la question « quoi ? » présente quelque chose de paradoxal. Le message, parlé ou écrit, est là ; à la question « quoi ? » on peut répondre, tautologiquement, par le message lui-même : tout y est. On peut à ce propos rappeler la réaction d'A. Breton qui, à un critique expliquant : « Saint Pol Roux a voulu dire que... », rétorquait « Si Saint Pol Roux avait voulu dire cela, vous pouvez être certain qu'il l'aurait dit ! ».

Il ne s'agit donc pas de paraphraser un texte, de le remplacer par un autre jugé plus clair, de faire de l'explication de textes. Et pourtant on note bien que la question « quoi ? », à propos d'un texte ou d'un discours, ou de toute autre forme de message, n'est pas vide de sens.

Au cours d'une enquête, puisque c'est de cela qu'il est principalement question ici, le chercheur, à une des étapes de son travail, se trouve confronté à un corpus constitué par les discours d'un certain nombre de personnes, toutes interrogées selon la même technique. On peut énumérer quelques-unes des questions qu'il est amené à se poser, au-delà du simple « qu'est-ce qu'il y a là-dedans ? » :

— Comment mettre chaque discours sous une forme plus maniable, de manière à en conserver tout ce qui est pertinent, et rien que cela ?

— Qu'a dit chacun à propos de tel point particulier ?

— Quelles différences et ressemblances y a-t-il entre les discours des personnes interrogées ?

Plus généralement, plus vaguement aussi, on souhaite mettre chaque discours sous une forme qui le rende plus « intelligible » avec tout ce que ce terme peut avoir de subjectif, d'ambigu, de flou. On veut obtenir un résumé [1] qui se présente sous une forme plus commode, par exemple pour pouvoir comparer plusieurs entretiens.

Tout ceci suppose une suite de jugements de commodité et de pertinence, donc une interaction de l'analyste qui s'appuie nécessairement sur des préalables étrangers au corpus lui-même, et qui dépend à la fois du cadre conceptuel dans lequel le chercheur travaille, et de ce qu'il veut faire du résultat de son travail. *Une analyse de contenu n'a de sens qu'orientée vers un objectif.* Chercher à savoir ce qu'il y a dans un texte, sans plus, n'a d'autre réponse que le texte lui-même.

Or ce n'est qu'en fonction de préalables très forts que l'analyste, par exemple, négligera ou non des fautes de syntaxe, admettra ou non que deux synonymes peuvent être considérés comme équivalents, que telle phrase entre ou non « dans le sujet », qu'il faut ou non tenir compte de l'ordre d'apparition des propositions, etc. Toute méthode d'analyse de contenu s'appuie sur de tels critères de pertinence, et on aura évidemment avantage à les expliciter.

---

1. Le mot « résumé » doit être pris ici dans son sens le plus large. Il n'est pas nécessairement plus court que le texte initial. Certaines techniques, par exemple, codent chaque mot selon plusieurs dimensions, ce qui augmente d'autant le volume du « résumé ». Celui-ci peut aussi ne pas être un texte au sens strict, mais peut se présenter, par exemple, sous la forme d'une liste de propositions, d'un graphe, d'un ensemble de diagrammes, de tableaux de fréquence, etc.

On pourrait, théoriquement, s'assigner un autre objectif, plus ambitieux : réduire le discours à un petit nombre d'éléments (propositions, relations) à partir desquels, par application de quelques règles de transformation, on pourrait le reconstituer dans son intégralité, comme on construit la totalité des expressions bien formées d'une algèbre à partir de ses générateurs et des opérations qui les transforment ; ou un ensemble de propositions à partir de postulats et de règles de déduction. C'est dans cette direction que s'orientent certaines recherches récentes qui visent à comprendre comment fonctionne « un texte ». Mais elles n'en sont pas encore aux applications pratiques, même si elles peuvent déjà inspirer certaines techniques d'analyse de contenu, elles ne permettent pas éventuellement d'éviter des jugements de pertinence et de commodité extérieurs au corpus étudié.

La seconde question qui peut être posée devant un ensemble d'entretiens, « qu'a dit chacun à propos de tel point particulier ? » est à première vue plus simple et plus évidente. Il s'agit simplement en apparence d'isoler dans chaque discours ce qui nous intéresse, cet intérêt étant, lui, explicitement intérieur au corpus. Mais cela pose des problèmes d'identification et de codage, donc de constitution de classes d'équivalences à l'intérieur desquelles on négligera des différences.

Quant à la troisième question qu'on peut reformuler comme « quel discours unique peut-on tenir sur l'ensemble des discours individuels recueillis ? », elle suppose, elle aussi, des critères d'équivalence et la systématisation des différences, et des hypothèses, extérieures au contenu lui-même, sur les conditions de production des différents discours.

L'importance du « quoi » ainsi spécifiée pour notre objet, nous conduira à négliger les autres termes de la question. En effet, on l'a vu, dans la pratique courante de l'analyse d'entretiens, le « qui a parlé », « à qui », est défini par la méthodologie de l'enquête ; quant à l'effet produit, il est en général négligeable dans la mesure où le seul but poursuivi par l'enquêteur est de « faire parler » l'enquêté. La question de l'effet produit serait, dans ce contexte, plus pertinente pour analyser l'effet des relances de l'enquêteur sur l'enquêté, et a été analysé par ailleurs (cf. 3.5.). Demeure la question du « comment », qui peut avoir quelque importance dans le cadre d'une enquête psychosociologique ou sociologique, mais qui dans les faits en a assez peu, dans la mesure où la préoccupation porte plus souvent sur *ce* qui est véhiculé que sur la forme utilisée pour le véhiculer [1].

6.4. MÉTHODOLOGIE DE L'ANALYSE DE CONTENU

La méthodologie générale de l'analyse de contenu répond essentiellement à deux types de questions : comment coder ? Comment s'assurer de la fiabilité de l'instrument ?

Ces questions sont elles-mêmes subdivisibles en un certain nombre d'autres que l'on

---

1. On ne reviendra pas ici, ni par la suite, sur le problème du sens, qui ne perd rien de son acuité, bien au contraire, mais qui a déjà été évoqué par ailleurs.

examinera. Auparavant, on aimerait rappeler que la question du codage est centrale dans la démarche de l'analyse de contenu, à double titre :

Tout d'abord pour des raisons théoriques que nous avons brièvement notées dans les pages précédentes, car au-delà du problème du codage, c'est le problème du sens qui est posé. On peut en effet penser que l'acte de coder est une opération sur le sens effectuée par le codeur, quelle que soit par ailleurs la technique de codage retenue. Ainsi, par exemple, coder les mots en fonction de leur fréquence d'apparition dans le texte peut paraître constituer une opération neutre. En fait, derrière cet acte existent un certain nombre d'hypothèses implicites ou explicites liées au statut accordé à la fréquence. Car, que dire sinon que tel mot apparaît n fois, tel autre mot n' fois, etc., et encore faudrait-il le dire de tous les mots ; en effet privilégier tel mot ou tel autre constitue déjà une opération sur le sens. On privilégie tel mot parce qu'il paraît « important », or l'importance est fonction d'une hypothèse sur le sens issu du statut de la fréquence (hypothèse qui pourrait être, entre autres, de type : une chose qui est dite souvent est importante pour celui qui la dit. Ce faisant, on postule implicitement que la fréquence d'apparition d'un mot connote celui-ci d'un sens particulier lié à l'émetteur). Mais, et c'est bien là le problème, il n'existe pas de théorie du sens qui puisse rendre compte de l'ensemble sémantico-situationnel.

Par conséquent, le seul recours possible, si l'on veut continuer à faire de l'analyse de contenu, est de se satisfaire d'une méthodologie du codage qui, n'ignorant pas les problèmes théoriques et posant nettement ses postulats, accepte le risque de l'empirisme. C'est-à-dire le risque d'une validation empirique et extrinsèque au champ d'analyse.

Ensuite, pour des raisons techniques, le codage doit répondre (et il ne le peut de façon absolue car il ne procède pas d'une théorie) aux trois critères d'objectivité, de systématicité, de généralité.

D'objectivité entendue dans le sens d'absence de liberté du codeur, ou encore de non-ambiguïté du code retenu. De systématicité et de généralité, dans la mesure où un code défini pour analyser une population d'entretiens donnés doit pouvoir être appliqué à l'ensemble de cette population d'une façon identique.

Ceci posé, il est certain qu'en l'absence du substrat théorique ces conditions ne peuvent être rigoureusement observées d'une part et que l'abus théorique interviendra au moment du traitement des résultats, c'est-à-dire au moment où l'on sera amené à dire « ce que j'ai trouvé *signifie* que... », d'autre part.

6.4.1. Comment coder ?

6.4.1.1. Catégories d'analyse

La première définition est celle de l'objectif de la recherche, transformé en catégories d'analyse [1]. On se posera la question de savoir si l'on s'intéresse aux valeurs, à l'origine, à la destination, etc.

1. On se place ici dans le cas de procédures closes. Dans le cas de procédures exploratoires, cette démarche ne s'impose pas.

TABLEAU 2

CATÉGORIES D'ANALYSE USUELLES

1. **Catégories « ce qui est dit »**

Matière du sujet	A quel propos est effectuée la communication ?
Direction	Comment est traitée la matière du sujet (par exemple favorable/défavorable, fort/faible) ?
Standard	Quelle est la base sur laquelle la classification par direction est faite ?
Valeurs	Quelles valeurs, buts, ou désirs, sont révélés ?
Méthodes	Quels moyens sont utilisés pour atteindre les buts ?
Traits	Quelles sont les caractéristiques utilisées dans la description de la personne ?
Autorité	Au nom de qui les énoncés sont-ils faits ?
Origine	Quelle est l'origine de la communication ?
Cible	Vers quelles personnes ou groupes la communication est dirigée ?
Localisation	Où l'action prend-t-elle place ?
Conflit	Quelles sont les sources et les leviers du conflit ?
Fins	Les conflits ont-ils été résolus de façon tragique, ambiguë, ou heureuse ?
Temps	Quand l'action prend-t-elle place ?

Mais aussi (et ceci n'est pas dû à Berelson) :

Etat	Quel est l'état de la communication à un moment donné ?
Code	Quel est le sous-code utilisé par le parleur ?

etc.

2. **Catégories « comment cela est dit »**

Forme ou type de la communication	Quel est le media utilisé pour communiquer (papier journal, radio, roman, télévision, etc.) ?
Forme des énoncés	Quelle est la forme grammaticale ou syntaxique de la communication ?
Dessein	Quelle est la méthode rhétorique ou de propagande utilisée ?

1. Dans le cadre d'une réponse au « Quoi ? » seules certaines catégories seraient à prendre en compte.

A ce propos, on peut reprendre la classification sommaire que Holsti (1968) rapporte à la suite de Berelson (1952) (tableau p. 165).

Berelson, à propos du choix des catégories, disait : « Les études... seront productives dans la mesure où les catégories seront clairement formulées et bien adaptées au problème et au contenu (à analyser) ». En fait, le tableau ci-dessus ne prétend nullement être exhaustif, tout au plus reflète-t-il les catégories les plus fréquemment utilisées.

Cette absence d'exhaustivité signifie qu'il ne saurait y avoir une standardisation de l'approche des problèmes et que chaque problème particulier peut relever d'un traitement spécifique. Par exemple, lors d'une étude sur les motivations d'étudiants de l'Ecole dentaire terminant leur scolarité, un traitement des interviews à partir des catégories « valeurs » et « type de communication » avait été réalisé. Un second traitement des réponses à des planches représentant graphiquement les valeurs isolées lors de la première analyse était dérivé de la méthode de dépouillement utilisée pour le T.A.T. La catégorie « ce qui est dit » n'entrait dans aucune des catégories définies dans le tableau présenté ci-dessus.

Ceci met l'accent sur l'impossibilité de standardisation et renvoie au caractère rigoureux de la démarche exigée par la définition des catégories d'analyse que l'on utilisera. C'est-à-dire, en fait, la façon dont on rendra opérationnel le problème à ce niveau. Non point que les tentatives pour standardiser l'approche d'un champ n'existent pas, mais elles se sont, pour la plupart, révélées inopérantes dès qu'on tentait de les généraliser à l'ensemble du champ.

Le problème issu de cette impossibilité provisoire est celui d'une non-comparabilité des résultats. Si, face à un problème spécifique, un chercheur ne peut faire appel à des catégories standards d'analyse, propres au champ dans lequel est situé le problème, si par conséquent il construit ses propres catégories, le risque est grand de ne pouvoir comparer les résultats obtenus par A sur un problème a, aux résultats obtenus par B sur un problème b, proche de a. En termes moins généraux, la comparaison est difficile, si un chercheur analyse les interactions de groupe avec la grille de Bales et un autre avec la grille de Mills. Difficulté importante, car elle est porteuse de désordres méthodologiques. Difficulté ayant suscité des tentatives pour unifier les catégories d'analyse dans tel ou tel champ particulier (psycholinguistique, motivationnel). Ce problème, insistons sur ce point, est lié à l'absence théorique. Il est tout à fait probable qu'une théorie des normes, par exemple, engendrerait un ensemble de concepts transformables en catégories d'analyse. En effet, quelle est la démarche conduisant à la définition de catégories ?

Le chercheur a un problème à résoudre, par exemple le repérage, dans des entretiens, des traits caractérisant l'attitude faciste. Il a, supposons-le, un *background* théorique (appelons-le ainsi pour les besoins de la cause, une autre cause nous conduirait à dire qu'une telle théorie est bien proche de l'empirisme) issu des travaux d'Adorno sur la personnalité autoritaire. Il va constituer ses catégories d'analyse conformément à ce *background*. Il isolera, par exemple, une catégorie rigidité/souplesse, une catégorie répressif/permissif, etc. Ensuite, il définira précisément les critères qui permettront de faire entrer telle unité d'analyse dans telle catégorie. En d'autres termes, il donnera un

sens précis (ou en tout cas le plus précis possible) à chacune de ses catégories. Différents moyens d'être « le plus précis possible » ont été utilisés, allant du recours à la théorie absente ou partielle à l'appel aux finesses méthodologiques ; l'une de ces dernières consistant à énumérer toutes les unités de contenu que les catégories devraient contenir, ce qui suppose, à tout le moins, une théorie solide, mais laquelle ? Ou une analyse préalable, mais sur quels critères ? Et on est ainsi renvoyé à une théorie solide.

Le problème est donc bien celui d'une absence théorique au niveau du sens (c'est-à-dire de la relation du discours et de sa production), face à laquelle on est amené à faire « au mieux ». Ce n'est pas là un constat pessimiste, mais l'analyse d'un état destinée à mettre en garde contre les limites de la méthode.

6.4.1.2. Unités d'enregistrement, de contexte, de numération

Après avoir effectué ce premier travail de définition des catégories, il convient d'analyser l'entretien. Pour ce faire, trois types d'unités sont généralement prises en compte :

— l'unité d'enregistrement,
— l'unité de contexte,
— l'unité de numération.

Pour comprendre les liens unissant ces trois types d'unités, on peut prendre un exemple simple. Supposons que les Halles de Baltard, au lieu d'être vouées à la démolition pure et simple, aient été démontées pour ensuite être remontées en un autre lieu. Il eut fallu analyser le contenu (métallique sans doute, mais un contenu tout de même) de celles-ci. Vraisemblablement, on aurait commencé par définir les unités d'enregistrement, c'est-à-dire, dans le cas présent, les pièces de charpente, puis on aurait défini les unités de contexte, c'est-à-dire l'ensemble nécessaire et suffisant pour comprendre le sens de la pièce de charpente (une voûte par exemple), enfin on aurait défini l'unité de numération (par exemple, le poids de chaque unité d'enregistrement).

Plus spécifiques à l'analyse d'entretien, si les unités d'enregistrement peuvent être définies comme le segment de contenu minimal qui est pris en compte par l'analyse (phrase, par exemple), l'unité de contexte est caractérisée par rapport à la précédente puisqu'elle est le plus étroit segment de contenu auquel on doit se référer pour comprendre l'unité d'enregistrement (paragraphe, par exemple). On pourrait également, dans un type donné d'analyse de contenu, avoir comme unité de contexte la phrase, et comme unité d'enregistrement le mot. Par contre, l'unité de numération est indépendante des deux précédentes, et peut ou non être utilisée concomitamment à l'unité d'enregistrement.

L'unité d'enregistrement peut être issue de plusieurs champs. On les classera de façon quelque peu arbitraire en : sémantique [1], linguistique et psychologique.

Chacun de ces types d'unité présente ses inconvénients, mais il semblerait que les unités linguistiques soient les moins suspectes quant à l'objectivité du codage, c'est-à-

1. La différence qui est établie entre unités « sémantiques » et « linguistiques » renvoie aux opérations effectuées par le codeur (soit celui-ci effectue des opérations sur le sens en codant, ou, si l'on préfère, il émet un métadiscours à l'occasion de la codification (unités sémantiques), soit se contente de découper le texte et d'enregistrer les unités lexicales isolées ou concaténées (unités linguistiques)).

dire quant à la réduction la plus complète possible du libre choix du codeur lors de l'opération de codage. Ce seront aussi celles qui permettront de dire le moins.

Parmi les unités sémantiques, reposant sur un choix issu d'opérations sur le sens, on trouve notamment l'une des unités d'enregistrement les plus utilisées : le thème. En effet, il y a bien opération sur le sens puisqu'une analyse thématique n'est rien d'autre que la reconnaissance de thèmes dans un discours qui n'est pas nécessairement thématique. Lorsqu'un codeur isole le thème « Liberté » dans une phrase telle que celle-ci « il rêvait souvent d'être parfaitement mobile, dégagé de toute entrave, de toute contrainte, un peu comme les chevaux blancs qui voilent la mer », il fait une opération sur le sens de cette phrase. Opération qui consiste à réécrire un fragment de discours pour en « tirer » le sens. Mais, rien à priori ne permet d'affirmer que c'est bien là le sens du texte, sinon « l'expérience » du codeur.

Une autre unité sémantique fréquemment utilisée est l'item, c'est-à-dire un entretien entier. Supposons que je veuille caractériser des entretiens en fonction de catégories préalablement définies : comique, triste, sentimental, bourgeois, etc., je choisirai comme unité d'enregistrement l'entretien, en d'autres termes l'item. Quoiqu'il en soit, cette unité demeure limitée dans son emploi à un certain nombre de problèmes spécifiques et ne va pas sans poser de nombreuses questions liées, notamment, à la taille macroscopique de l'unité considérée. En effet, où classer un entretien triste et comique (ce qui peut arriver si celui-ci a deux parties par exemple) ?

Parmi les unités linguistiques, le mot a été souvent employé et l'ordinateur a fait croître ce courant d'utilisation. En général, le mot est utilisé à des fins de décompte fréquentiel (combien de fois les policiers interrogés enploient-ils le mot « voyous ») destiné à caractériser un ensemble de choses différentes : les schizophrènes, le style d'Alain Decaux, les entretiens « au coin du feu », la possibilité d'attribuer le *Roman de la Momie* à Marguerite Duras, etc.

Par ailleurs, un ensemble de mots particuliers, grammaticalement caractérisés, renvoie soit à des hypothèses théoriques explicites et génératrices de catégories, par exemple utilisation des pronoms personnels liée à une attitude linguistique de nature syntagmatique, soit au contraire à une pure exploration : on a remarqué que les schizophrènes émettaient davantage de substantifs, moins de verbes, etc.

Une autre unité d'enregistrement peut être la période, caractérisée comme étant la phrase (Vignaux). On trouvera également le morphème (Harris), défini comme un élément isolable, et on trouvera encore la proposition. Cette liste ne se veut pas exhaustive, elle met simplement l'accent sur des méthodologies quelque peu négligées par les psychosociologues. On peut cependant penser que des interrogations des problèmatiques, qui sous-tendent le choix de telle ou telle unité, doivent être effectuées. On le verra dans la partie suivante.

L'unité de contexte est, ainsi que nous l'avons dit, le plus étroit segment de contenu nécessaire pour comprendre l'unité d'enregistrement. Il n'y a pas grand-chose de plus à en dire, sinon que le choix de l'unité de contexte risque de modifier les résultats de l'analyse. Une expérience de Geller, Kaplan et Laswell, comparant quatre ensembles d'unités d'enregistrement et d'unités de contexte appliquées au même contenu, montre que si les quatre méthodes indiquent bien le biais idéologique des journaux analysés de

la même façon, elles diffèrent de façon importante quant à l'extension du biais considéré. De façon générale, il semblerait que plus l'unité d'enregistrement est importante, plus le biais croît, d'une part, et plus la taille de l'unité de contexte croît plus le nombre d'entrées neutres [1] diminue, d'autre part. Osgood — mais nous citons cela à titre d'indication seulement — suggère que la taille d'une unité de contexte pourrait être située entre 120 et 210 mots. Une telle prise de position suppose résolu un tel nombre de problèmes sémantiques qu'il semble difficile d'accorder une valeur de norme à ces limites.

L'unité de numération peut être de deux sortes : arithmétique ou géométrique.

Arithmétique, elle servira à compter le nombre de fois où apparaît tel contenu, le temps d'antenne consacré à tel sujet, le nombre d'images représentant tel type de situation, le temps consacré à l'accomplissement de tel type de tâche, l'intensité des attitudes (échelles), etc. Géométrique, elle servira à mesurer l'espace consacré à tel contenu. Seule nous intéresse la première sorte, dans le cadre de l'analyse d'entretiens.

Il faut être attentif au fait que le choix d'un système de numération n'est pas neutre. Le simple fait d'accomplir une opération quantitative signifie déjà que l'on accorde un statut à l'instrument de quantification. Ainsi relever la fréquence d'apparition d'un mot donné peut signifier que l'on considère que plus ce mot apparaît fréquemment dans le discours d'un émetteur, plus ce qu'il signifie a d'importance pour l'émetteur. Mais cette affirmation est difficilement soutenable dans certains cas et il est nécessaire d'introduire d'autres mesures telles que celle de l'intensité. Par conséquent, le choix d'un système de numération quelconque doit donner lieu à l'explicitation des présupposés qu'il engendre, car ici comme précédemment, et bien que cela soit moins évident au premier abord, se posera le problème du sens à double titre. Si l'on reprend l'exemple précédent, le statut de la fréquence confère au mot un sens dû au nombre d'apparitions, sens qui vient connoter le sens usuel du terme. Autrement dit, le fait de postuler que la fréquence d'apparition d'un terme est significative de l'importance du concept qu'il exprime, ajoute au sens du terme. Il y a là opération sur le sens. Mais il y a également opération cognitive. On postule, effectuant cette opération, qu'il existe chez l'émetteur un certain type de comportement. Ce dernier peut être caractérisé, entre autres, par la répétition des choses considérées comme importantes. Et si c'était l'inverse, si l'émetteur ne parlait jamais des choses importantes, quel moyen a-t-on de vérifier une absence (dans ce cadre-là tout au moins) ? Ceci n'est qu'un exemple pour montrer qu'un grand nombre de problèmes théoriques peuvent se poser à partir d'une opération apparemment simple.

Quant aux problèmes méthodologiques posés par la numération, on peut en trouver un exemple dans l'association du relevé fréquentiel et de la mesure de l'intensité. Supposons qu'au cours d'un entretien soient prononcées les cinq phrases suivantes :

1. Je ne peux vraiment pas supporter Alexandre.
2. Alexandre est entré dans la pièce et je me suis mis à le haïr.
3. Bien sûr, il y a des moments de répit dans ma maladie pendant lesquels mes relations avec Alexandre s'améliorent.

---

1. Les biais étaient caractérisés comme attitude favorable, neutre, défavorable.

4. J'aimerais voir Alexandre pendu.
5. Alexandre était si bon avec moi, que parfois je me prenais à l'aimer.

Si, voulant analyser l'attitude du sujet à l'égard d'Alexandre, je me contente de relever le nombre de fois où Alexandre est cité, il est évident que je ne pourrais en tirer aucune conclusion valable. Dans des cas de ce genre, qui relèvent d'interrogations méthodologiques, on fait souvent intervenir la mesure de l'intensité. On citera, à titre d'exemple, la méthode de comparaisons par paires (voir 7.2.1.2.6.). Celle-ci est bâtie sur la méthode des juges qui utilisent une échelle d'attitude linéaire pour classer des paires d'indicateurs. Les jugements servent ensuite à constituer les catégories dans lesquelles seront classées les unités d'enregistrement.

Ainsi, le problème de la numération est celui d'un double obstacle : théorique et méthodologique. Mal le poser théoriquement revient à émettre des conclusions non fondées, mal le poser méthodologiquement conduit à utiliser des outils inadéquats.

6.4.2. Fiabilité de l'instrument ?

Cette interrogation renvoie à trois types de problèmes communs aux sciences sociales :

— l'échantillonnage des données,
— la fidélité,
— la validité.

Pour ce qui est du premier terme, il ne diffère pas de ce qui a été dit à propos de l'échantillonnage en général (cf. chapitre 2) ; on analysera par contre plus longuement les deux autres termes caractérisant la fiabilité de l'instrument.

6.4.2.1. Fidélité

La fidélité de l'instrument est liée à la procédure de codage, et par conséquent au codeur et à l'instrument de codage dont il dispose. En fonction de cela, les tests de fidélité porteront essentiellement sur les deux pôles que sont la fidélité du codeur et la fidélité des catégories d'analyse. La fidélité du codeur doit être envisagée sous deux aspects : inter codeur et intra codeur.

Soit, dans le premier cas, un ensemble de codeurs opérant sur un même texte doit aboutir aux mêmes résultats, et, dans le deuxième cas, un même codeur analysant le même texte à deux moments différents doit reproduire sa première analyse. Ceci revient à dire que les résultats doivent être indépendants de ceux qui les produisent. Une absence de fidélité peut cependant avoir deux causes que l'on examinera à la fin de ce paragraphe : le codeur ou les catégories qu'il utilise.

Le deuxième problème auquel se heurte la fidélité est celui des catégories d'analyse utilisées par le codeur. Le maître mot est ici ambiguïté. Dans la mesure où la catégorie utilisée n'est pas ambiguë, c'est-à-dire permet de classer sans difficulté l'unité d'enregistrement, la fidélité est complète. Cependant, au-delà de la catégorie, l'unité d'enregistrement elle-même peut être source de désaccords, le cas est fréquent lorsqu'il s'agit d'analyses thématiques.

Les moyens de réduire les risques attachés à telle ou telle partie du processus sont limités et leur garantie est seulement méthodologique. En ce qui concerne le codeur, il semble qu'il n'y ait guère qu'une préparation suffisante qui puisse permettre d'atteindre à la fidélité souhaitée. Cependant, le travail du codeur n'est pas indépendant de l'outil dont il dispose et la non-ambiguïté des catégories, ainsi que celle des unités d'enregistrement, permet une fidélité plus grande. En effet, si les catégories et/ou les unités d'enregistrement sont ambiguës (ce qui est le cas de bon nombre de catégories et de quelques types d'unités d'enregistrement, dont la plus employée est le thème) que signifie avoir un taux de fidélité élevé pour un codeur ? Quel que soit le cas de figure (inter ou intra) le taux de fidélité ne peut être élevé que si le codeur a bien intégré les normes du chercheur. Ceci ne signifie donc pas que le codeur est techniquement valable, cela signifie tout au plus qu'il est idéologiquement façonné.

Par conséquent, moins que la formation du codeur, pour assurer la fidélité, il semble que le problème doive être d'abord de se préoccuper des catégories d'analyse et des unités d'enregistrement. Si celles-ci sont parfaitement non ambiguës, le reste n'est plus que l'acquisition d'un savoir-faire. Le manque de fidélité dans ce cas relevant d'une ignorance et non d'une déficience méthodologique.

Tout le problème consistera donc à rendre les catégories claires et les unités d'enregistrement strictement définies. De nombreuses méthodes, allant de l'utilisation des jugements catégoriels dichotomiques à l'utilisation de juges additionnels, ont été employées, mais ce ne sont là que des palliatifs méthodologiques à une absence théorique. Absence d'autant plus cruelle dans ce cas qu'elle est le moyen de la validité de l'outil.

Indépendamment de ces problèmes, ou plus exactement de façon indépendante, ont été développés un certain nombre de coefficients destinés à mesurer la fidélité des codeurs ou des catégories. On renverra le lecteur aux indices présentés dans R. Daval et coll. (1964). Disons simplement que ces indices sont de trois types : fidélité directe des codeurs, fidélité directe des catégories et fidélité moyenne et pondérée des catégories. En fait, seuls les troisième et quatrième indices présentent quelque intérêt dans la mesure où, contrairement aux deux précédents, ils s'intéressent à la relation existant entre les catégories et les codeurs. Les deux autres indices testent, le premier la fidélité inter codeur, le second « la fidélité reconnue à chaque catégorie par le groupe tout entier » ; seul les deux suivants permettent de réintroduire les poids individuels dans la cotation.

On demeure cependant persuadé que le problème est moins dans la mise au point d'indices que dans la définition rigoureuse des catégories et des unités d'enregistrement qui les peupleront. Or c'est surtout un problème théorique ainsi que cela a été souligné, et non, dans la plupart des cas tout au moins, méthodologique.

6.4.2.2. Validité

La validité pourrait être définie comme l'adéquation entre les buts et les fins, sans distorsion des faits. D'où la nécessaire prise en compte des différentes étapes qui ont été décrites jusqu'à présent : représentativité de l'échantillon, choix des catégories, indices de quantification, etc.

Tout au plus peut-on, pour aller un peu plus loin dans l'analyse du concept de validité, considérer que celui-ci est très lié aux buts poursuivis. Ainsi l'*American Psychological Association Committee on Psychological Tests* établit une distinction entre quatre sortes de validité : validité du contenu, validité prédictive, validité comparative, validité interprétative. Chacun de ces types de validité est lié à des recherches poursuivant des buts différents, mais ne renvoyant pas nécessairement à l'utilisation de techniques différentes.

Afin d'éclairer ce qui précède, on peut prendre un exemple emprunté à des travaux sur les alcooliques. Une batterie de ratios syntaxiques permet de rendre compte de phrases différentes dans le vécu de l'alcoolique. Ces ratios sont bâtis à partir de relevés fréquentiels de mots (verbes, adjectifs, noms, etc.). On pourrait vouloir mesurer le vocabulaire des alcooliques et, dans ce cas, on s'intéresserait à la validité du contenu. On pourrait également essayer de prédire la durée de la cure de désintoxication et on serait alors renvoyé à la validité prédictive. On pourrait encore comparer le vocabulaire des alcooliques et des non-alcooliques, il s'agirait de validité comparative. On pourrait enfin souhaiter faire des inférences sur les capacités mnémoniques des sujets alcooliques, on serait dans le cas de validité interprétative.

Ainsi qu'on le voit, la validité n'est rien d'autre qu'une notion méthodologique et un concept lié aux buts poursuivis par une recherche.

6.5. QUELQUES MÉTHODES D'ANALYSE DE CONTENU

Il ne s'agit pas ici de rendre compte de l'ensemble des méthodes d'analyse de contenu existant et ce pour trois raisons. La première est qu'un ouvrage entier n'y suffirait pas, à moins de lui donner la forme caractéristique des *Trois Mousquetaires*, suivi de *Quelque Temps après*. La seconde, plus importante, réside dans la qualité des méthodes d'analyse de contenu proposées. En effet, si certaines, malgré les problèmes théoriques qu'elles posent, se révèlent d'une fiabilité pratique importante, un grand nombre ne présentent qu'un faible intérêt. Non point qu'elles soient méprisables mais elles ont des adeptes et des défenseurs autrement plus compétents que les auteurs de cet ouvrage. La troisième raison, enfin, tient à l'objet qu'on s'est proposé concernant l'analyse de contenu, et qui conduit à éliminer un certain nombre d'analyses dites du discours dont on ne voit pas précisément quel pourrait être leur intérêt dans le cadre d'une enquête psychosociologique. Le choix fait est donc, nécessairement, partiel ou partial, et ne prend pas en compte le mépris profond dans lequel les linguistes tiennent l'analyse de contenu pratiquée par les psychosociologues. Mépris justifié par une absence théorique qu'aucune pratique linguistique ne peut malheureusement combler à ce jour. Et pourtant combien on aimerait que les conseilleurs soient ici également les fournisseurs !

6.5.1. Procédure close

La distinction entre procédures close et ouverte renvoie on l'a vu à deux grands types d'analyse de contenu : celles qui font intervenir des catégories définies préalablement à l'analyse, et celles qui n'en font pas intervenir, ayant par là même un caractère purement exploratoire.

Dans ce paragraphe, on s'intéressera donc à quelques méthodes d'analyse de contenu procédant de catégories préalablement fixées. Il faut entendre par-là que l'analyse est liée à un cadre catégoriel empirique et/ou théorique qui le sous-tend et auquel elle se réfère.

Cinq types de modalités catégorielles seront pris en compte : psychologique, psycholinguistique, psychosociologique, linguistique, documentaire.

6.5.1.1. Catégories psychologiques

L'utilisation du terme « psychologique » pour désigner des catégories renvoie à la psychologie individuelle. Ceci signifie que le cadre théorique de référence est un cadre de psychologie individuelle tirant le plus souvent son origine d'une source clinique. La grille de lecture des tests de personnalité constitue le lieu privilégié dans lequel on peut puiser afin de créer des catégories utilisables dans d'autres domaines.

On présentera deux recherches qui permettront de décrire la méthode tout en l'illustrant par une application.

La première avait pour objet d'étudier l'image que de futurs praticiens dentistes se faisaient d'eux-mêmes et de leur relation avec le patient. Elle concernait des étudiants de cinquième année [1] de l'Ecole dentaire de Paris (Dangy, Ghiglione, Beauvois, 1973).

Après avoir analysé de façon thématique des entretiens menés auprès d'eux, il apparaissait qu'il était nécessaire de dépasser le niveau des représentations et de l'idéologie que l'on avait atteint, afin de préciser comment les étudiants de cinquième année vivaient et se situaient par rapport aux principaux aspects de la relation praticien/patient. Ces principaux aspects avaient été déterminés à partir de l'analyse thématique d'entretiens et de l'analyse de concomitances thématiques.

Pour atteindre le but poursuivi dans cette phase il importait que le matériel utilisé réponde à deux exigences :

— Être suffisamment ambigu pour permettre certaines projections et plusieurs types de défenses.

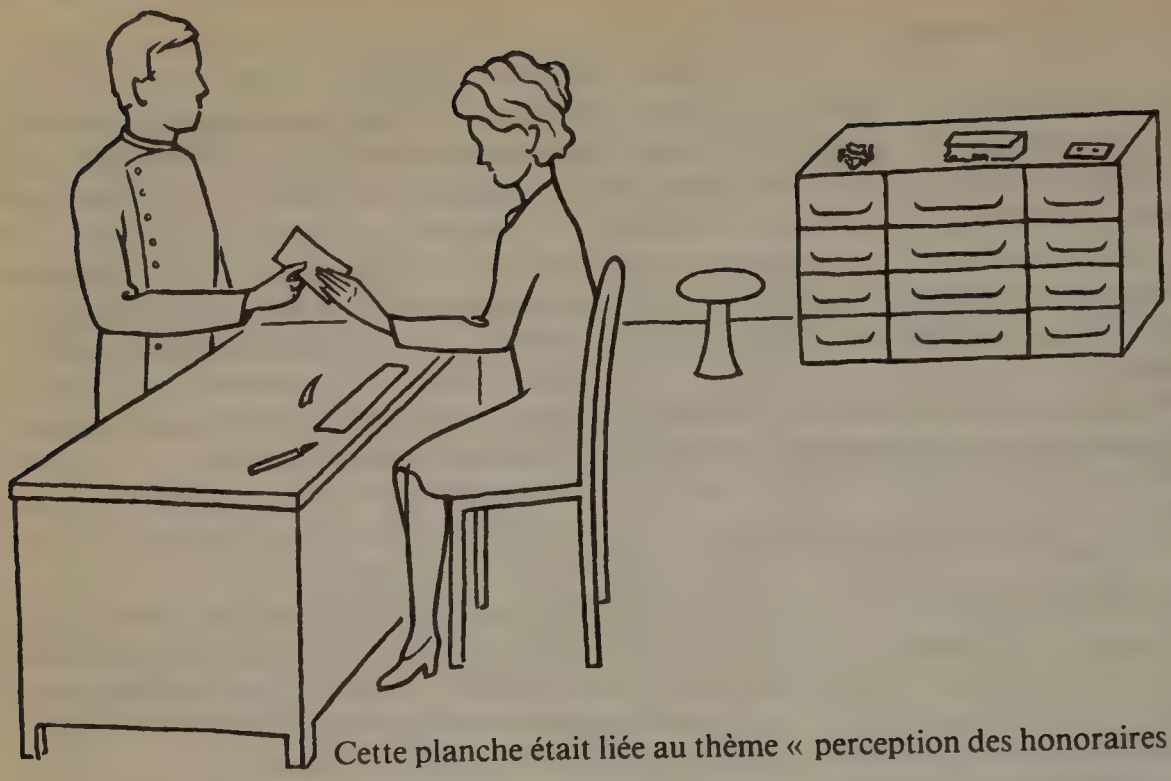
— Être suffisamment représentatif des divers problèmes qui devaient être étudiés.

L'utilisation de tests projectifs ayant été envisagée, il a semblé qu'une méthodologie proche de celle du T.A.T. répondait aux exigences formulées.

Cinq planches ont été dessinées représentant de façon ambiguë les cinq problèmes clés de la relation praticien/patient. Une grille d'analyse des protocoles, issue des principes des tests à intention projective, a été mise au point. A titre d'exemple on présentera ci-dessous la planche 3 et l'interprétation à laquelle elle a donné lieu.

---

1. C'est-à-dire la dernière année d'études. Les élèves peuvent exercer dès la fin de celle-ci.



Cette planche était liée au thème « perception des honoraires ».

« Il a déjà été discuté de ce problème à la lumière des entretiens. Nous retrouvons ici les principaux éléments signalés, jusqu'à la garantie idéologique médicale puisque c'est pour cette planche seule que, quelquefois, le médecin est substitué au dentiste. Un certain nombre d'indices signalent le caractère problématique de la planche. Le paiement des honoraires est perçu d'emblée 11 fois sur 27. L'ambiguïté du matériel n'est pas en cause puisque effectivement à quatre reprises des sujets attribuent d'abord une autre signification à la planche (feuille de sécurité sociale plutôt que billet de banque), puis on arrive au problème des honoraires après coup pour finalement dire des choses semblables à celles dites par les onze sujets précédents. Enfin, dans deux cas le paiement des honoraires perçu d'emblée est par la suite oublié à la faveur d'un changement de signification attribué au matériel.

C'est à cette planche que la situation de test est la plus refusée, de nombreux sujets ne parvenant pas à raconter une histoire transforment la situation en situation d'entretien afin de pouvoir produire un discours cohérent et idéologiquement rationalisé. D'ailleurs à cette planche apparaissent :

— des effets discursifs particulièrement fréquents,
— des annulations,
— des dénégations.

Les dénégations observées sont généralement représentatives de l'incompatibilité entre diverses sphères idéologiques qui apparaissent simultanément dans le discours. Cette incompatibilité peut produire des attitudes chez le praticien dentiste décrites ainsi « il est gêné, mais obligé », « il est embarrassé », « il a mauvaise conscience », des conduites d'évitements magiques ou par délégation... »

A travers la trame interprétative on voit donc apparaître une catégorisation préalable (dénégation, annulation, perception de la signification, etc.) qui relève d'une rationalité théorique extérieure et qui donne son sens au texte.

La seconde recherche porte sur un entretien avec Gérard Zwang « La sensualité ambiguë » reproduit dans *Elle*. La méthode d'analyse consistait à rechercher à l'intérieur du texte les thèmes ressortant des principes de réalité et de plaisir (Freud). La lecture du texte, et partant son sens, étaient déterminés par une théorisation externe, la psychanalyse.

L'analyse conduisait aux résultats que l'on trouvera p. 176-177.

Ces résultats tendaient à confirmer l'hypothèse selon laquelle il ne s'agit pas dans l'article d'une argumentation fondée sur une théorie psychanalytique rigoureuse, mais sur un « psychanalysme » de bon aloi. C'est-à-dire en fait que les arguments n'ont pas d'appui théorique, seul le reflet déformé d'une théorie en tient place.

6.5.1.2. Catégories psycholinguistiques

Différentes recherches ont tenté — et tentent, la source ne s'étant pas tarie — d'étudier les traits psychologiques des individus à partir de leurs écrits ou de leurs discours.

A titre d'exemple on citera brièvement une recherche menée sur les schizophrènes en 1971, à l'Université de Paris VIII-Vincennes.

L'hypothèse de départ était ainsi formulée : « Le langage aliéné reflétant les structures de la maladie mentale, on peut retrouver les caractéristiques du langage d'un schizophrène dans des fragments de discours d'un sujet malade. »

Un *background* conceptuel existait déjà dans le domaine et permettait de définir les caractéristiques d'un langage émis par les schizophrènes. Ainsi pour Racamier « la sémantique et la syntaxe y sont sacrifiées au profit du rythme et de la phonétique », « le vocabulaire se restreint dans son répertoire. Certains mots sont expurgés, des autres se répètent ». Pour Ey on peut également noter que :

— La conversation est parfois impossible (mutisme ou semi-mutisme), coupée d'implusions verbales (injures, blasphèmes, obscénités) ou de réponses à côté. De plus, celle-ci est toujours singulière car elle n'est pas destinée à établir un contact.

— Au point de vue phonétique, intonation, rythme, articulation, sont désintégrés. Il en va de même pour la structure des mots dont l'unité est rompue.

— Au point de vue sémantique le signifié est détourné de son accord avec le langage commun, il y a perte de la référence au code ou perte de la contiguité. Il y a création de néologismes, incohérence (salade de mots, galimatias, langage hermétique).

Enfin, d'après divers auteurs américains :

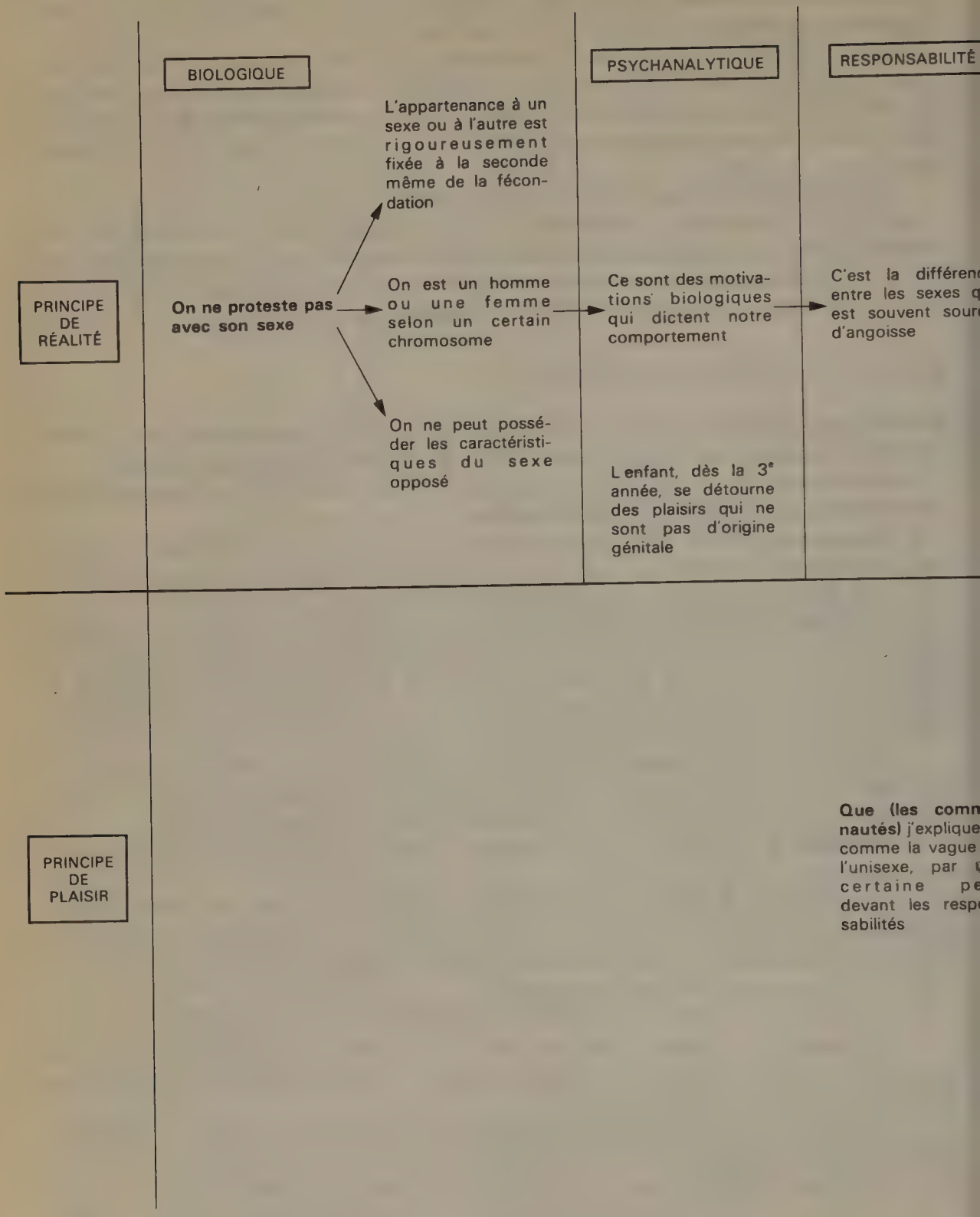
— Au point de vue syntaxique différentes altérations peuvent être relevées.

— Le Type Token Ratio (TTR) qui mesure le nombre de mots différents dans des ensembles de 100, 200, 500 ou 1 000 mots, montre que la moyenne des mots différents est plus faible chez les schizophrènes.

— Le rapport « adjectif-verbe » qui mesure le nombre d'adjectifs pour 100 verbes montre que les sujets normaux ont un rapport plus élevé que les schizophrènes.

— Le rapport « adjectif-nom » va dans le même sens que le précédent.

A partir de ces différentes considérations et d'une méthodologie utilisée par Gonin


BIOLOGIQUE
PSYCHANALYTIQUE
RESPONSABILITÉ
PRINCIPE DE RÉALITÉ
On ne proteste pas avec son sexe
L'appartenance à un sexe ou à l'autre est rigoureusement fixée à la seconde même de la fécondation
On est un homme ou une femme selon un certain chromosome
On ne peut posséder les caractéristiques du sexe opposé
Ce sont des motivations biologiques qui dictent notre comportement
L enfant, dès la 3e année, se détourne des plaisirs qui ne sont pas d'origine génitale
C'est la différenc entre les sexes q est souvent sour d'angoisse
PRINCIPE DE PLAISIR
Que (les comm nautés) j'explique comme la vague l'unisexe, par t certaine pe devant les resp sabilités

ut que chacun
ure à l'autre
uvissement de
ésir, et par des
ens qui sont
es à son sexe

Un homme se retrouve face à une femme, et l'amour n'est pas une formalité entre copains

Chacun doit mener son propre combat

passe d'un
na rigide à un
conformisme
emble surtout
re la peur de
nation de son
e sexe

---

oiller de la
façon, vivre
nde, cela per-
'abolir la diffé-
donc d'avoir
u moins peur
utre

nalheureux de
leur mère,
nt de devenir
t de la com-
uté

une source de
ité rassurante
moment

Les communautés sont peuplées de gens dont le caractère est souvent un peu faible

Comme la peur devant la différence sexuelle explique le comportement unisexe, elle explique la transexualité

Il existe des hétérosexuels que la femme féminine excite beaucoup

Ils ont, pour parler en termes freudiens, un très fort narcissisme qui provoque chez eux l'envie de se parer

Le plus souvent, ce sont des hommes qui veulent échapper à leur devoir de virilité

Et peut-être se faire opérer pour être délivré du besoin d'avoir à prouver leur virilité

(1969) il a été mis au point une procédure d'analyse. A partir d'un fragment de discours correspondant à mille lexies, définies par Gonin ainsi « toutes les fois qu'une suite de mots est figée par l'usage et possède un sens global qui dessine le contour d'un concept nouveau par rapport aux mots de cette suite, cette suite est considérée comme une seule partie du discours ou lexie, exemple : point de vue, chef de gare », on a procédé à un calcul de la fréquence des classes grammaticales. Par ailleurs les principaux rapports définis par les auteurs américains ont donné lieu à analyse. Les résultats obtenus sont présentés dans les tableaux suivants.

Une analyse qualitative menée parallèlement a permis également de mettre en évidence différents traits du langage du sujet schizophrène :

1. D'un point de vue psychiatrique, on peut relever une impression générale d'incohérence, un rythme du discours désordonné, l'apparition de thèmes marquant le morcellement de la personnalité (problème d'identité, indistinction du locuteur et du locuté), le passage d'un discours adapté à un discours délirant.

2. D'un point de vue linguistique, une étude du tableau des résultats montre une nette diminution des substantifs et adjectifs. Par ailleurs, ce vocabulaire pauvre est surtout constitué de paraphrases verbales, de paraphrases littérales, de néologismes. Enfin, une étude de la grammaire met en évidence une dissyntaxie. En effet, il y a augmentation des mots d'articulation (en particulier des conjonctions de subordination et de coordination) et mauvaise utilisation de ceux-ci.

6.5.1.3. Catégories psychosociologiques

Elles peuvent être divisées en deux rubriques selon qu'elle tentent d'analyser des textes en les référant à ce que les Américains nomment des « standards » ou qu'elles s'intéressent à l'étude des interactions de groupe.

Les « standards ». Ce type de recherche fait appel à des concepts issus de théorisations psychosociales ou d'études empiriques.

En ce qui concerne les théorisations psychosociales, celles qui produisent des « standard » relèvent généralement du domaine des attitudes. Utiliser le mot de théorie est peut-être ambitieux quand on voit l'état de celle-ci et les recherches effectuées dans ce domaine. Ces dernières ont pour la plupart été bâties sur des catégories attitudinelles empiriques, dont le moins qu'on puisse dire est qu'elles sont contestables.

Quoi qu'il en soit, un certain nombre de recherches réfèrent à des typologies connues du type *need of achievement* et *need of affiliation*. Ainsi des études consacrées aux changements culturels. On citera, à titre d'exemple, l'étude de Mc Clelland. Selon cet auteur un individu présentant un haut degré d'*achievement* « est quelqu'un qui désire réussir, qui est énergique et non conforme et qui aime les tâches présentant un risque ». L'*achievement* a été défini opérationnellement comme « une somme du nombre d'exemples, d'idées ou images d'*achievement* ». L'hypothèse selon laquelle « une société avec un pourcentage élevé d'individus ayant un haut *n achievement* contiendra une classe entrepreneuriale forte qui tendra à être active et réussira particulièrement dans les entreprises commerciales et industrielles de telle sorte que la société globale croîtra en pouvoir et influence », fut testée en examinant la littérature grecque durant trois périodes : croissance (900-475 av. J.-C.), stabilité (475-362 av. J.-C.), déclin (362-

CLASSES GRAMMATICALES	NOMBRE BRUT		POURCENTAGE		VALEUR DU X2	SEUIL DE SIGNIFICATION			OBSERVATIONS
	Témoins	Schizophrène	Témoins	Schizophrène					
LEXIES.....	1000	1000							
MOTS	1080	1080							
SUBSTANTIFS ...	203	146	18	13	11,1			.001	en nette diminution
ADJECTIFS.....	72	50	7	4,6	4,2	.01		.05	baisse sensible
ADVERBES.....	61	75	5	7	1,5		.10		équivalence
VERBES :									
Principaux	75	86	7	8	0,8		.10		équivalence
Auxiliaires	63	60	5	5	0,7		.10		équivalence
A l'infinitif	17	32	2	3	4,7	.01		.05	sensible augmentation
PRONOMS.....	186	260	17	22	37,3			.001	nette augmentation
Je	35	77	3	7	25,4			.001	nette augmentation
Tu - vous.....	4	36	0,3	3	26			.001	nette augmentation
Il - elle.....	12	7	1	0,6	1,3		.10		équivalence
PRÉPOSITIONS...	102	118	9	10	1,3		.10		équivalence
CONJONCTIONS .	50	78	4,6	7	6,5	.01		.05	sensible augmentation
ARTICLES.....	114	58	10,1	5	19,7			.001	baisse sensible
INTERJECTIONS..	13	13	1	1					
NÉGATIONS.....	27	49	2	4	6,6	.01		.05	sensible augmentation
NÉOLOGISMES...	0	13		1					

	TÉMOINS	SCHIZOPHRÈNES
« Adjectif - verbe »	45 %	25 %
« Adjectif - nom »	38 %	34 %

100 av. J.-C.). Par ailleurs un indicateur du pouvoir économique et social grec et de sa zone d'influence fut trouvé dans la localisation géographique des vases grecs découverts. Ceci permettait de mesurer en unités métriques la zone d'échanges commerciaux et d'influence des Grecs. Entre le sixième et le quatrième siècle avant J.-C. celle-ci passe de 1,2 à 1,9 million de miles carrés. Le tableau ci-dessous présente les résultats d'exemples d'*achievement* trouvés dans la littérature (d'après Holsti, 1968) :

Nombre d'exemples de *n achievement* pour 100 lignes par type d'échantillon.

Période	L'homme et ses dieux	Direction de l'État	Cérémonie funéraire	Poésie	Epigram - mes	Discours de guerre	Moyenne
900-475 av. J.C.	2,01	3,54	7,93	2,87	4,72	7,38	4,74
475-362 av. J.C.	1,21	0,82	5,94	0,38	2,36	5,55	2,71
362-100 av. J.C.	0,81	0,00	2,54	0,16	1,57	3,00	1,35

D'autres études confirment ces résultats, bien que le matériel et la méthodologie utilisés soient légèrement différents. Une autre dimension attitudinelle fort utilisée a été celle d'introversion/extraversion, liée suivant les circonstances aux besoins de réalisation de soi (*achievement)* ou d'affiliation au groupe (*affiliation).*

Pour ce qui est des concepts empiriques, la littérature en fournit également de nombreux exemples. Trois méthodes de recherche semblent toutefois être plus fréquemment utilisées et donner des résultats empiriques plus satisfaisants dans l'ensemble que ceux généralement obtenus par ailleurs. La première renvoie à la constitution de valeurs de référence à l'aide d'une enquête préliminaire dans le domaine étudié. Ce qui signifie qu'à la différence des exemples précédents ces valeurs sont propres à chaque problème étudié ou à chaque famille de problèmes, et ne peuvent pas être généralisées — comme c'est le cas des concepts de *need of achievement/need of affiliation* ou introversion/extraversion — à un ensemble de problèmes très divers. Par exemple, une recherche portant sur les valeurs de l'automobiliste était fondée sur les concepts suivants, concepts déterminés au cours d'une enquête préliminaire (voir tableau p. 181).

On notera toute l'ambiguïté de certains termes caractérisant les valeurs, et par là même la difficulté que l'analyste peut parfois éprouver pour classer un thème. Par exemple, les notions de juste ou de normal sont-elles plus précisément psychosociales ou morales ? Les auteurs de la recherche ont cru bon de placer des notions sous la rubrique « valeurs morales », mais on peut se demander en fonction de quels critères ils ont opéré.

Une autre recherche présentant moins de risques méthodologiques, mais procédant

VALEURS	CARACTÉRISATION
1. Pratiques	Sécurité Rapidité Facilité Confort Agrément Economie Utilité etc.
2. Biologiques	Vie Santé Satisfaction des besoins physiques etc.
3. Psychosociologiques et psychosociales	Prestige Réussite Moderne A la mode etc.
4. Morales	Bien Juste Normal etc.

de la même démarche, est celle de Larson, Gray et Fortis (1963) tentant d'étudier les buts poursuivis et les résultats obtenus, tels qu'ils étaient décrits dans les programmes de télévision pour enfants, adultes, et enfants/adultes. Les « standards » utilisés étaient « socialement approuvé (légal) », ou socialement désapprouvé (violent, non légal). C'est-à-dire, en fait, des standards bâtis sur des normes sociales. La méthode consistait à coder chaque acte dont on pouvait clairement indiquer la relation avec un but particulier. Paradoxalement (?) les résultats ont montré que les actes ayant les plus faibles chances de succès dans l'atteinte des buts étaient ceux qui étaient définis comme socialement approuvés. Par ailleurs, il n'apparaissait à cet égard aucune différence entre les programmes pour enfants et les programmes pour adultes.

En fait, un des risques majeurs de la méthode réside dans l'empirisme qui a présidé à sa construction. Dans le cas précédent, le repérage dans un texte des exemples de besoin

de réalisation de soi *(need of achievement)*, une fois celui-ci opérationnellement défini, présente un risque moindre car il s'agit de décider si tel ou tel fait rapporté est conforme à la définition opérationnelle qui en est faite. On aurait presque tendance à dire que le phénomène du sens est dans ce cas secondaire car la recherche est liée à des considérations extra-linguistiques au plan théorique, et le repérage est également effectué à partir des critères extra-linguistiques. Le problème n'est plus, dès lors, sémantique mais psychosocial. La théorie critiquable est extérieure au langage et seule la méthode de repérage peut éventuellement (mais tout dépend de l'outil utilisé) donner lieu à un débat sémantique, prenant la forme suivante : est-ce que cet exemple (son sens) est bien propre à illustrer le « besoin de réalisation de soi » ? Le problème réside alors dans l'affinement de l'outil, une fois la théorie qui le sous-tend admise. Dans le deuxième cas, il n'en va pas de même. On travaille toujours au niveau langagier, sans aucun support théorique. Les catégories standard procèdent le plus souvent d'une analyse thématique, ce qui suppose donc des opérations sur le sens, leurs spécifications également. Quant à l'analyse proprement dite, à moins de se contenter d'un repérage des mots spécifiant les valeurs, elle ne peut, à son tour, que procéder par opérations sur le sens.

Une seconde catégorie de méthodes est liée à l'utilisation de facteurs de situation ou de circonstances :

— Qui, avec qui, pour qui ? (facteurs individuels).
— Où, quand ? (facteurs qualitatifs/quantitatifs).
— Pourquoi ? (facteurs motivationnels).

Cette méthode présentant moins de risques que la précédente, car purement descriptive, a été (et est vraisemblablement encore dans la société de consommation où se situe la rationalité de la demande méthodologique) utilisée, notamment pour étudier les habitudes d'achat des « ménagères ». Pour prendre rapidement un exemple, non par honte mais par crainte de lasser, un responsable de supermarché peut poser des questions et obtenir des réponses du genre de celles que nous trouvons dans le tableau de la p. 183 (pure imagination d'un auteur se délassant).

D'où l'astucieuse stratégie du commerçant imaginant les circuits de progression dans les rayons soigneusement étudiés (non Madame, vous ne pourrez aller au rayon alimentation sans passer obligatoirement par les rayons disques, vêtements, bijoux, etc., et être soumise aux tentations qui assaillent la pauvre ménagère), la présentation sur les rayons, particulièrement calculée, le renforcement musical, olfactif, visuel, étonnamment gradué et le consommateur remarquablement piégé.

Le troisième type de méthodes regroupe toutes les méthodologies issues de l'utilisation d'un savoir social antérieur : dans cette rubrique on trouve l'étude des thèmes exclus ou minimisés, des stéréotypes apparaissant dans le discours, des fragments de messages « sociaux » (publicitaires, dus à l'enquêteur, à des articles de journaux, etc.). En fait, cette approche pourrait être caractérisée comme étant la recherche des traces dans un discours donné. Elle présente un grand nombre d'avantages lorsqu'on tente d'analyser les modalités d'influence d'un discours sur un auditeur en prenant en compte les traces mnésiques laissées par ce discours. Toutefois, le problème méthodologique lié à cette approche est celui du repérage des traces. On peut penser en effet que, sauf

Qui vient ?	Les femmes d'ouvriers spécialisés (40 %), d'ETAM (40 %), de cadres moyens (20 %)
Avec qui ?	Avec leur mère (20 %), leurs enfants (70 %), seules (10 %)
Pour qui ?	Pour elles et/ou leur famille (90 %), pour leur chien (10 %)
Où vont-elles dans le magasin ?	60 % au rayon alimentation seul, 20 % aux rayons : alimentation et produits d'entretien, 10 % aux rayons : alimentation, produits d'entretien, fournitures pour la maison, 10 % au rayon produits de beauté seul
Quand viennent-elles ?	Le lundi (25 % de l'effectif hebdomadaire), le mercredi 25 %, le samedi 30 % etc.

cas exceptionnels, ces traces ne sont pas restituées telles quelles dans le discours du sujet, ce qui conduit à effectuer des opérations sur le sens n'ayant pas leur contrepartie théorique.

6.5.1.4. Catégories linguistiques

Ces dernières années, les recherches en ce domaine se sont fortement développées, vraisemblablement parce que l'analyse de contenu, « rebaptisée » analyse du discours, trouvait là des théorisations linguistiques partielles, mais existantes, sur lesquelles elle pouvait fonder sa pratique.

Cependant il semble qu'en ce domaine aussi on soit également confronté aux abus de sens. Entendons par-là à une utilisation non fondée du sémantique qui, lui, n'a pas produit de théorie satisfaisante. Autrement dit, l'impression qui se dégage des travaux récents tendrait à laisser penser que l'on a trouvé une nouvelle langue descriptive dont le formalisme est plus ou moins justifié théoriquement et méthodologiquement, mais que l'on n'a pu aller plus loin. De telle façon que la transformation d'un message en un autre, à l'aide d'un ensemble d'opérations formalisées, semble être la seule voie possible. Déborder conduit quasi inéluctablement à commettre des abus de pouvoir, à faire dire aux résultats plus qu'ils ne peuvent.

Harris, parlant de sa méthode d'analyse distributionnelle, dit qu'elle est capable de déterminer *comment* le texte dit, non *ce* qu'il dit. Pour Harris, en effet, l'« analyse distributionnelle permet d'obtenir des résultats précis sur certaines relations dépassant les limites de la phrase si nous ne considérons que les phrases d'un seul discours suivi,

c'est-à-dire celles qui ont été prononcées ou écrites les unes à la suite des autres par une ou plusieurs personnes dans une seule situation ». Ce qui ne permet aucune opération sur le sens et, par là même, ne présente pas un grand intérêt pour les objectifs que poursuit l'analyse de contenu.

Les méthodes dérivées des travaux d'Harris sont nombreuses, mais toutes sont limitées au comment. Dès qu'elles s'en écartent, on pense aux analyses d'Achard et de Vignaux, elles deviennent fragiles car elles effectuent des opérations sur le sens ; quant à Pecheux, dont les ambitions sont vastes, il avoue toutefois « qu'un écart... sépare l'analyse du discours de la théorie du discours ».

Autrement dit, le choix des linguistes risque d'enfermer l'analyse du discours dans l'univers linguistique, la condamnant à ne pouvoir lier le linguistique à l'extra-linguistique. Les quelques tentatives formalisées de liaison ne sont guère convaincantes dans la mesure où elles échouent au point crucial, la théorisation de la liaison entre le linguistique et l'extra-linguistique.

6.5.2. Procédures ouvertes

Ce type d'analyse porte également dans la littérature le terme de « procédure exploratoire » (Henry et Moscivici, 1968), c'est-à-dire qu'aucun cadre catégoriel théorique ou empirique ne sert de support à l'analyse. Les résultats sont dus à la seule méthodologie d'analyse exempte de tout appel à un cadre préétabli. En d'autres termes on pourrait dire que la méthodologie est auto-génératrice du sens du résultat. Ceci ne va pas évidemment sans un grand nombre d'aspects criticables, dont celui concernant le sens du « dit » n'est pas le moindre.

On ne citera pas plus ici que précédemment l'ensemble des méthodes existantes dans le cadre des procédures exploratoires [1].

6.5.2.1. Le décompte fréquentiel

Le but des décomptes fréquentiels, dans le cadre ici fixé, peut être défini comme l'établissement d'« une sorte d'inventaire des diverses opinions, ou des arguments utilisés », mais aussi des mots, etc. En somme l'important est ce que Pinto et Grawitz, nomment « une sorte d'inventaire », car cela signifie que l'information nouvelle apportée par l'analyse de contenu eu égard au texte analysé réside dans la constatation de la fréquence ou dans sa comparaison avec d'autres fréquences. Ceci pose le problème crucial du statut de la fréquence. En effet, dire que parce que tel journal présente plus fréquemment des thèmes favorables à tel parti il est proche de ce parti, peut s'avérer parfois extrêmement périlleux. Cette affirmation conduit à donner un sens à la fréquence d'apparition des thèmes sans que celui-ci ne soit justifié par rien d'autre que par l'expérience phénoménologique du chercheur.

---

1. Il faut noter que ces méthodes peuvent être également utilisées dans les procédures closes (ainsi de l'analyse thématique, du décompte fréquentiel, etc.). Si on les a reportées dans cette partie c'est tout simplement parce qu'elles sont plus fréquemment utilisées en tant que telles et non comme outils dans une procédure ouverte d'une part, et qu'il était plus commode, pour des raisons de clarté, de les regrouper d'autre part.

6.5.2.2. L'ANALYSE THÉMATIQUE

De loin la plus fréquemment utilisée, elle est aussi une de celle qui pose le plus de problèmes car elle combine des opérations sur le sens avec des opérations méthodologiques hasardeuses.

Sa pratique consiste à isoler des thèmes dans un texte afin de le ramener à des proportions utilisables d'une part, et à permettre sa comparaison avec d'autres textes traités de la même façon d'autre part. Pour entrer un peu plus avant dans la méthode, on établit généralement une distinction entre les thèmes principaux et les thèmes secondaires. Les premiers peuvent être définis comme rendant compte du contenu du segment de texte analysé, les seconds servant à spécifier le premier dans ses différents aspects. Par exemple :

THÈME PRINCIPAL	THÈME SECONDAIRE	TEXTE
Sécurité	• Détente • Tranquillité d'esprit • Opposition sécurité/vitesse	I. Vous comprenez quand je suis dans ma bagnole ce qui compte le plus c'est d'être tranquille question danger. Moi je veux pas avoir de crainte, je veux me sentir à l'aise, détendu, l'esprit tranquille. Il me faut une voiture sûre. Alors vous comprenez moi les petites voitures — je les appelle les danseuses — qui font vroom et qui vont vite mais ne vous garantissent pas, ça je m'en tape.

Mais, lorsqu'un naïf demande comment on a pu arriver à un tel résultat, la réponse aussi élaborée soit-elle ne peut masquer le fait que personne n'en sait rien, y compris celui qui l'a atteint. Hormis les phrases du type « mais c'est évident enfin », ou leurs dérivées, personne ne sait définir un thème dans son extension (d'où les nécessaires unités de contexte) ni caractériser les opérations sur le sens qui ont conduit à « sortir » le thème « sécurité ». Bien sûr il y a le vécu de l'analyste et son expérience du langage et de la méthode, mais il n'y a que ça et ce n'est pas suffisant. D'ailleurs, à analyser l'analyse thématique, on en arrive à la conclusion que seule l'expérience du codeur peut remplacer la double défaillance théorique et méthodologique (ce qui est un record car les analyses pèchent souvent par le premier type de défaillance, mais évitent assez bien le second). Et pourtant... 90 % des analyses dites de motivation (pourquoi Paic, Madame ?) utilisent cette méthode et bâtissent de « mirobolantes » publicités sur les thèmes qui sont les plus fréquents dans le discours. Et pourtant... l'analyse thématique n'est qu'un métadiscours émis par le codeur sur le discours et personne ne peut analyser les

conditions de production de ce métadiscours et son fonctionnement. Alors autant demander au codeur de parler sur le discours des autres. Ce sera plus simple, moins onéreux et presque aussi efficace.

6.5.2.3. Les concomitances thématiques

Cette méthode à visée structuraliste consiste à repérer dans la linéarité, et avec des contraintes de distance, les thèmes qui apparaissent, dans un entretien ou un ensemble d'entretiens, les uns avec les autres ou, au contraire, qui n'apparaissent jamais les uns avec les autres. Cette méthode ne résout pas les problèmes cités au-dessus mais apporte parfois des suppléments d'information dans la mesure où on compare des fréquences d'apparition théoriques (en faisant comme si la répartition des thèmes était homogène au sein d'un ensemble de discours) dans un corpus avec les fréquences d'apparition réelles. En reprenant l'exemple de l'étude « Praticien-patient », une telle analyse, dont on trouvera ci-après le tableau, conduit à conclure qu'il existe dans le corpus analysé deux types d'idéologies : l'une structurée autour des préoccupations de l'étudiant, l'autre structurée à partir des problèmes supposés du praticien exerçant (voir p. 187).

Ce type d'analyse est au croisement des deux méthodes que constituent l'analyse thématique et le décompte fréquentiel.

6.5.2.4. L'analyse par grappes

Cette méthode, mise au point et expérimentée (Ghiglione et Matalon, 1974) souffre d'un certain nombre d'imperfections. Son souci était de répondre à la défaillance constatée au niveau de l'analyse thématique, et non à la défaillance sémantique, car cette dernière relève d'une autre approche du problème.

La méthode

Elle comporte cinq étapes :

a) On fait l'hypothèse structuraliste selon laquelle il existe dans un texte [1] (et plus particulièrement dans celui qu'on analyse) un nombre limité de référents noyaux intervenant dans le discours, principalement comme sujets des propositions constitutives du discours. Ces référents noyaux sont considérés comme autant de paradigmes que l'on peut désigner par un substantif ou un pronom personnel, et ils sont constitués d'un terme générique auquel peut se substituer un ensemble d'équivalents paradigmatiques (pronoms personnels, substantifs, pronoms relatifs, articles définis, pronoms possessifs, adverbes).

La première étape consiste à isoler les référents noyaux en tenant compte de deux exigences :

— l'ensemble des référents noyaux isolés doit rendre compte du plus grand nombre de propositions possible,

— les référents noyaux ne doivent pas être trop importants en nombre.

1. Et plus particulièrement dans une interview, où notre méthode permet d'éliminer une redondance que l'on sait être forte.

TABLEAU DES CONCOMITANCES

Fré-quence	Fréquence	16	22	10	10	12	12	12	13	15	14	10	15	9
	Thèmes / Thèmes	3.2	7.3	1.3	1.4	1.1	4.1	11.2	10.3	6.1	9.3	14	6.6	6.7
16	3.2			14 [11,60]	7 [5,1]	8 [5.1]	7 [6.3]	3 [6,3]	5 [6,3]	7 [6,6]	7 [7,8]	8 [7,2]	5 [5,1]	6 [7,8]	5 [4,8]
22	7.3				9 [7,2]	8 [7,2]	9 [9]	8 [5]	6 [9]	11 [9,3]	12 [10,8]	12 [10,2]	7 [7,2]	10 [10,8]	7 [6,5]
10	1.3					5 [3]	5 [3,9]	3 [3,9]	4 [3,9]	6 [4,2]	6 [4,8]	6 [4,5]	3 [3]	5 [1,8]	3 [2,8]
10	1.4						5 [3,9]	4 [3,9]	3 [3,9]	3 [4,2]	6 [4,8]	7 [4,5]	3 [3]	4 [4,8]	1 [4,8]
12	1.1							5 [4,8]	7 [4,8]	7 [5,1]	7 [6]	5 [5,4]	2 [3,9]	2 [6]	1 [3,6]
12	4.1								4 [4,8]	7 [5,1]	8 [6]	5 [5,4]	3 [3,9]	5 [6]	3 [3,6]
12	11.2									5 [5,1]	5 [6]	4 [5,4]	6 [3,9]	2 [6]	2 [3,6]
13	10.3										6 [6,3]	6 [6]	4 [4,2]	6 [6,4]	5 [3,3]
15	6.1											6 [7]	5 [4,3]	7 [7,5]	3 [4,5]
14	9.3												6 [4,5]	8 [7]	8 [4,2]
10	14													8 [4,8]	5 [2,8]
15	6.6														6 [4,5]
9	6.7														

Les chiffres non entourés, dans les cases, représentent les fréquences réelles d'apparition de chaque concomitance.
Les chiffres entourés représentent les concomitances théoriques.

b) Ayant créé une liste de référents noyaux, l'analyste doit réécrire le texte en ne retenant que les propositions (définies comme un ensemble comprenant un sujet/un verbe/un syntagme complément) contenant le référent noyau ou son équivalent. Si le fragment de discours est structuré en proposition principale et subordonnée, on note entre parenthèses la subordonnée [1] afin de ne pas courir le risque de rendre difficilement compréhensible la proposition principale. Cette réécriture aboutit à un ensemble de grappes de propositions segmentées par référent noyau.

c) Après cette réécriture on procède à des réductions à l'intérieur de chaque grappe :

— Par élimination des équivalences synonymiques, définies soit comme deux propositions identiques, soit comme deux propositions contenant le même syntagme complément. Par exemple : le psychologue doit être à l'écoute de l'autre, et : le rôle du psychologue est d'être à l'écoute de l'autre.

— Par élimination des propositions conséquentes impliquées dans une proposition antécédente :

— le premier cas d'implication est celui où le sujet d'une proposition (P1) est inclus dans le sujet d'une autre proposition (P2). On conserve P2. Par exemple :

P2. Tous les étudiants pensent que le psychologue est utile,

P1. Les étudiants de ce cours pensent que le psychologue est utile,

— le second cas d'implication est celui où le sujet est le même et où le prédicat d'une proposition (P1) est inclus dans le prédicat de l'autre (P2). Par exemple :

P2. Tous les étudiants pensent que le psychologue est socialement utile,

P1. Tous les étudiants pensent que le psychologue est utile.

— Par ailleurs peuvent être faites des réductions par implication ou par synonymie fondées sur une connaissance préalable de ces relations par le chercheur, attribuée également au sujet. Par exemple :

P2. Les psychologues expérimentalistes disent que la nuit les rats blancs sont gris,

P1. Le professeur N... dit que la nuit les rats blancs sont gris.

Or, il est de notoriété publique que le professeur N... est un psychologue expérimentaliste, donc P1 est éliminé.

A cette phase de réduction peuvent être adjointes d'autres réductions provenant d'hypothèses émises par le chercheur. Par exemple : celui-ci n'est pas intéressé par l'histoire personnelle du sujet qui parle et décide d'éliminer toutes les propositions contenant des indications liées à l'histoire du sujet et que personne ne pourrait donner à sa place, bien qu'appartenant au même groupe institutionnel, etc. Ou encore il met l'accent sur l'acteur et élimine, si deux propositions sont situées dans deux grappes différentes, celle des propositions où le référent noyau est complément.

d) Après qu'aient été opérées les diverses réductions, la quatrième phase est de restructuration :

— A l'intérieur de la même grappe sont regroupées par genre toutes les propositions faisant intervenir la même spécification du référent. Par exemple le référent noyau : psychologue, verra l'ensemble des propositions contenant le référent noyau : psychologue-pathologue, regroupé dans la même classe.

1. C'est là qu'éventuellement un système d'équivalence de type Harris pourrait s'avérer utile à notre propos.

— Puis entre grappes est opérée une restructuration intertextuelle à partir des syntagmes compléments caractérisant le référent noyau. Cette dernière restructuration conduit à définir des équivalences entre les référents noyaux.

e) Enfin, la dernière étape est interprétative et on sait pertinemment qu'une théorie sémantique la rendrait plus crédible, car dans l'état actuel des connaissances celle-ci ne peut qu'être attribuée à l'intuition du chercheur.

A titre d'exemple, on présentera le résultat d'une analyse effectuée selon cette méthode.

a) Référents noyaux :

1.1 Le(s) psychologue(s)
1.2 La psychologie
1.3 Le(s) malade(s)

b) Découpage du texte en propositions autour des référents noyaux :

2.1 *Le(s) psychologue(s)*

P1. *Gens* [1] qui essaient de comprendre les hommes.
P2. *Qui* étudient principalement leurs fonctions intellectuelles, affectives.
P3. *Ils* ont commencé par étudier l'enfant (pour arriver ensuite à l'adulte).
P4. *Ils* veulent aider l'homme de différentes manières.
P5. Les professions de *psychologue,* c'est aider les hommes (quand ils ont des problèmes affectifs ou intellectuels).
P6. Les *psychologues* qui s'intéressent à l'orientation professionnelle font passer des tests intellectuels.
P7. Le pathologue, c'est un *psychologue.*
P8. La pathologie est différente d'un *psychologue* (qui travaille uniquement sur la pédagogie).
P9. Un *psychologue* dans l'industrie.
P10. Un *psychologue* dans l'éducation.
P11. Sur telle ou telle catégorie de *psychologue,* je serais plus apte à répondre.
P12. Le *psychopathologue* (je le vois comme) quelqu'un qui travaille dans les hôpitaux psychiatriques.
P13. Le *psychologue,* il doit toujours être à l'écoute des autres.
P14. *Son* rôle essentiel, c'est d'écouter les autres.
P15. *Il* en déduit ce qu'il veut après.
P16. *Il* a un rôle très important (dans la mesure où).
P17. *Il* peut aider plus facilement un malade mental.
P18. *Son* rôle, c'est d'être à son écoute, essayer de le comprendre.
P19. Le *psychologue,* donc le pathologue, il est là pour écouter.
P20. Le *psychopathologue* est là pour essayer de faire progresser le malade.
P21. (Je vois pas exactement) le rôle du *psychologue* dans l'industrie.

1. Les mots soulignés sont les équivalents du mot psychologue.

P22. (Discuter avec les travailleurs sur les conditions de travail) je ne sais pas si c'est exactement son rôle à *lui*.

P23. Il y a le *psychologue* orienteur professionnel.

P24. *Il* se contente de faire passer des tests.

P25. (Des éducateurs spécialisés) qui sont des *psychologues*.

P26. L'*éducateur spécialisé*, son problème, c'est l'éducation d'enfants retardés ou caractériels ou débiles mentaux.

P27. Là aussi, il y a une grande part d'écoute de la part du *psychologue*.

P28. Ça rejoint un peu le rôle du *psychopathologue*.

P29. Les *psychologues* généticiens, ils sont spécialement tournés vers l'enfance, vers la petite enfance.

P30. Pour *eux*, il y a beaucoup moins ce rôle d'écoute.

2.2 *La psychologie*

P'1. (les étudiants en *psychologie)*, ils ont des idées sur leur avenir qui sont très différentes.

P'2. Je suis plus renseignée sur la *psychologie* de l'éducation (que dans le domaine industriel par exemple).

2.3 *Le malade*

P''1. Il peut plus facilement aider un *malade mental* (mieux essayer de le comprendre).

P''2. Le psychopathologue est là pour essayer de faire progresser le *malade*.

P''3. En essayant d'analyser les causes de *sa* maladie.

P''4. En essayant d'arriver à le faire sortir de *sa* maladie.

P''5. Un *malade*, ce qu'il veut avant tout c'est qu'on l'écoute.

P''6. *Il* a besoin de raconter sa vie ou ses ennuis.

c) Réduction 1 :

Sont exclues les propositions :

P9. Car le syntagme « psychologue dans l'industrie » se retrouve en P21, où l'information apportée est supérieure.

P23. et P24. Car ces deux propositions sont, prises conjointement, synonymes de P6 :

P6. Les psychologues qui s'intéressent à l'orientation professionnelle font passer des tests intellectuels.

P23. et P24. Le psychologue orienteur professionnel... il se contente de faire passer des tests.

Réduction 2 :

Sont exclues les propositions :

P2. — P4. qui sont incluses dans la proposition P5.

P7. qui est incluse dans P12. (syntagme psychopathologue).

P13. incluse dans P14.

P18. incluse dans P19.

P22. incluse dans P21.

Réduction 3 :

P''1. qu'on retrouve également en P17.
P''2. qu'on retrouve également en P20.

d) Réduction relevant des hypothèses du chercheur :
Choix de supprimer tout ce qui est propre au sujet et non généralisable à sa classe institutionnelle d'appartenance (ici : les étudiants).
P11. :...*je* serais plus apte à répondre.
P21. : *je* ne vois pas exactement...
P'2. : *je* suis plus renseignée...

e) Regroupement des propositions par nature :
— A l'intérieur de la grappe :
6.1. Le *psychologue*
6.1.1. *Le psychologue en général*
P1.P3.P5.P14.P15.
6.1.2. *Le psychopathologue*
P12.P16.P17.P18.P20.P28.
6.1.3. *Le psychologue éducateur spécialisé*
P25.P26.P27.P28.
6.1.4. *Autres psychologues*
P6.P10.P29.P30.
6.2. *La psychologie*
P'1.
6.3. *Le malade*
P''3. P''4. P''5. P''6.
— Entre grappes :
— Le psychopathologue est à regrouper avec le malade.
— L'éducateur spécialisé est à regrouper avec le psychopathologue.
— Le psychologue en général est à regrouper avec le psychopathologue (voir le tableau p. 192).

f) Interprétation et conclusions :
— Le psychologue en général est assimilé au psychopathologue et ce à travers les caractéristiques communes :
— d'écoute,
— d'aide,
— de compréhension.
— L'éducateur spécialisé est assimilé au psychologue (à travers des propositions identiques pour le pathologue et l'éducateur : des éducateurs spécialisés qui sont des psychologues ; le pathologue, c'est un psychologue) et au psychopathologue par le canal de l'écoute.
— Le « malade mental », le « malade », l'« homme » ont des demandes symétriques au rôle dévolu au psychologue.
— Les autres psychologues (et les étudiants) sont ignorés ou définis par un stéréotype.

L'ÉDUCATEUR SPÉCIALISÉ	LE PSYCHOLOGUE EN GÉNÉRAL	LE PATHOLOGUE
P27. Là aussi, il y a une grande part d'écoute de la part du psychologue	P1. « essaie de comprendre les hommes »	P17. « peut aider plus facilement »
	P5. « aider les hommes »	P18. « son rôle, c'est d'être à son écoute, essayer de le comprendre »
	P14. « écouter les autres »	P20. « essayer de faire progresser le malade »
		Le malade
P28. Ça rejoint un peu le rôle du psychopathologue		P"5. « un malade il veut avant tout qu'on l'écoute » P"6. « il a besoin de raconter »

NOTA : Le nombre de propositions (5) qui leur est consacré par rapport aux propositions définissant le psychologue et le pathologue et l'éducateur spécialisé (14) est intéressant à noter.

Soit une représentation du type :

— Le psychologue — psychopathologue : écoute, aide, comprend.

— Le malade veut qu'on l'écoute.

— Les autres psychologues (industriels, généticiens, orienteurs professionnels) sont dévalorisés.

6.5.2.5. L'ANALYSE PAR CHAMPS SÉMANTIQUES

Cette analyse beaucoup moins complexe que la précédente a pour objet d'analyser les caractéristiques attachées à une entité donnée. Les principes méthodologiques, simplifiés, sont les suivants :

a) définition du ou des référents noyaux,

b) relevé de tous les termes caractérisant le référent noyau et étant situés dans une phrase contenant à la fois le référent noyau, ou son équivalent, et le terme qualifiant ce dernier,

c) réduction en fonction des synonymies, ou des inclusions (ses yeux bleus, son œil bleu),

d) structuration des termes, intratextuelle, et des référents, intertextuelle,

e) interprétation.

A titre d'exemple, on citera une étude dont le but était de définir le portrait de héros de roman d'espionnage dans ses aspects physiques et psychologiques. Les romans pris en compte concernaient neuf héros différents. La méthodologie a conduit à ne retenir, pour caractériser le héros défini a priori comme étant le seul référent noyau que :

— les adjectifs qualificatifs (épithètes ou attributs),
— les noms auxquels est adjoint un adjectif, en apposition,
— les périphrases désignant le héros (une « sorte de machine efficace »).

Par contre, ont été rejetés :

— les termes trop directement liés à l'action et non au sujet (exemple : il est responsable de la mort de...),
— les syntagmes verbaux (exemple : redoutant d'être aperçu),
— les adverbes (exemple : remercia chaudement),
— les formes ambiguës, même introduites par le verbe (« libre à vous d'avoir une opinion »).

Le tableau ci-après donne un aperçu des résultats de l'analyse en ce qui concerne les caractéristiques des héros analysés (p. 194 sq.).

6.6. L'AGRÉGATION DES ANALYSES

Au cours des pages qui précèdent nous avons parlé de l'analyse de contenu d'un texte ou d'un discours uniques, éventuellement d'un corpus considéré comme homogène. Lorsqu'on analyse un ensemble d'entretiens cette homogénéité n'est certainement pas assurée, comme on pourrait dans certains cas admettre qu'elle l'est, dans les différents articles d'un même journal, par exemple. Or, en général, on ne peut se contenter de la simple juxtaposition des analyses des entretiens individuels : il est nécessaire d'en faire une synthèse, c'est-à-dire de tenir sur l'ensemble de ceux-ci un discours unique.

Il arrive qu'on puisse faire l'économie de ce problème. Par exemple, lorsqu'on cherche à savoir tout ce qui se dit à propos d'un certain thème, pour ensuite construire un questionnaire. On se contentera alors d'une simple énumération de propositions, sans s'inquiéter de savoir qui a dit quoi, si certaines phrases ont été dites par la même personne ou par des personnes différentes ; on se bornera au plus à noter la fréquence de chaque énoncé. Mais les cas où on peut se contenter de résultats aussi sommaires sont rares.

Souvent le problème est escamoté : on traite, sans justification explicite, l'ensemble des entretiens comme un texte unique. Cela donne des affirmations comme celle-ci, que nous avons rencontrée dans un rapport d'enquête d'un bureau d'études important : « la population étudiée est ambivalente à l'égard de...», sans qu'il soit possible de savoir si l'analyse a montré que les individus eux-mêmes, ou au moins la majorité d'entre eux, sont ambivalents, ou si une partie de la population est favorable et l'autre défavorable. Il est évident que les conclusions qu'on pourra tirer de l'une ou l'autre interprétation ont toutes les chances de ne pas être les mêmes, il est donc important de les distinguer de façon rigoureuse.

Dans les analyses de contenu thématiques classiques on essaie de traiter ce problème en distinguant les analyses « verticales » et « horizontales ». L'analyse verticale, c'est

	OSS 117	SUZUKI	KERN	LE COMMANDER
ASPECT GÉNÉRAL	Poitrine large et dure	Une espèce de bonze au crâne rasé		Haute taille Blond
ALLURE AIR	Félin	Allure inquiétante d'un clergyman, manières cassantes et sournoises, sérieux		Air goguenard
VISAGE	Lèvres pleines	Traits ascétiques, spiritualité et douceur du sourire Détermination implacable	Sourire guoguenard, cruel Moue excédée	Teint hâlé Sourire amusé
YEUX REGARD	Yeux clairs Regard circonspect	Yeux légèrement bridés Regard impérieux	Regard goguenard	Regard clair, dur, inquiétant, insoutenable
VOIX TON	Voix grave et tranquille, ton sans réplique	Voix dure et métallique, timbre grave et vibrant, ton calme	Voix théâtrale	
ORIGINE FONCTION	Américain Hubert Bonisseur de la Bath	Agent de la CIA Bouddhiste Zen	Intelligence Officier de la CIA Colonel, un « pro » de l'entourloupe	Américain Agent de l'Office of Navy Intelligence
HABILLEMENT	Pantalon tergal, polo, semelles de caoutchouc			

DIVERS	Bras protecteur et rassurant	Gourou		
INTELLECTUELLE	Excellente mémoire visuelle		Lucide Sceptique	Très fort (= intelligent), sceptique
INSTINCT	Sens aigu de l'orientation, réflexes bien conditionnés			
ATTENTION CONTRÔLE VOLONTÉ	Volonté de fer	Imperturbable Implacable	Impassible Flegmatique Pince sans rire	Intransigeant
COMPORTEMENT	Toujours prudent, la précision même, tempérament combatif	Tranquille		Prudent
AFFECTION	Impitoyable Sentimental quelquefois (= plus humain)			Maladivement méfiant
SENSATION ÉMOTION			Agacé Mal à l'aise	
MORALE PHILOSOPHIE	Horreur de tuer sans raison impérieuse	Bouddhiste Zen	Puritain	
SEXUALITÉ	Pas l'homme d'une seule femme			
DIVERS	Enigmatique	Impénétrable	Etrange	Impénétrable

	S.A.S.	COPLAN	GAUNCE	MATT	FACE D'ANGE
ASPECT GÉNÉRAL	Grand blond très séduisant	Grand gaillard	Une sorte de machine	La trentaine 1,80 m 80 kgs de muscles	Minceur musclée Blondeur lisse et dorée Beauté froide
ALLURE AIR	Maintien nonchalant		Allure souple Air suave		Allure féline Air soucieux, moqueur
VISAGE	Mince sourire Sourire grinçant	Masque viril Moue perplexe	Rude visage bronzé Rides Mèche rebelle	Triangulaire Petites rides Demi-sourire Cheveux noirs et courts striés de gris	
YEUX REGARD	Yeux d'un fauve Yeux d'or Lunettes	Neutre Inquisiteur, lueur meurtrière	Yeux d'un gris très pâle, yeux brillants Mystérieux, déterminé, expression angoissée	Yeux clairs bleus virant au vert	Pâleur minérale, yeux bleus Eclat froid, œil attentif
VOIX TON	Voix froide Ton sec	Voix contenue, sèche Ton catégorique	Péremptoire	Très mâle, dangereusement séduisant	Voix calme et froide

	S.A.S.	COPLAN	GAUNCE	MAT-T	FACE D'ANGE
ORIGINE FONCTION	Autrichien, Son Altesse Sérénissime le Prince Malko Linge Barbouze de luxe (CIA)	Agent français	Directeur européen des services de la CIA	Ancêtres : les hurons, américain Agent de valeur	Dernier preux chevalier en exercice
HABILLEMENT	Impeccablement vêtu, décorations	Elégant	Sale et barbu		
DIVERS	Magnifique demeure historique en Autriche	Main persuasive	Marié	Main ferme et professionnelle	
INTELLECTUELLE	Cerveau en forme d'IBM. Remarquable homme d'affaires	Intuition Machiavélisme	Homme intelligent Imagination		Intelligent
INSTINCT		Instantanément sur la défensive		Flair et instinct de conservation	
ATTENTION CONTRÔLE VOLONTÉ	Tendu détendu	Equilibre intérieur, stable, imperturbable, attentif, tendu détendu, confiance en lui-même	Homme d'acier Intransigeant Impassible L'œil partout	Détermination Sûr de réussir	

	S.A.S.	COPLAN	GAUNCE	MATT	FACE D'ANGE
COMPORTEMENT	Gentleman	Fair play, menaçant, impulsif, cordial, narquois, dynamisme féroce	Un fonceur, brutalité, bagarreur, homme tranquille, sardonique	Brutalité Aucune intention agressive	
AFFECTION	Euphorique, rage folle, haine froide, désespéré	Sentimental	Désabusé Indifférent Soucieux	Pas systématiquement méfiant	Sentimental
SENSATION ÉMOTION	Vexé surpris Pas surpris, bouleversé, dégoûté, en supplice, pris de panique	Exaspéré	Sensibilité épidermique, épouvanté, ému, agacé		
MORALE PHILOSOPHIE	Très fataliste, horreur de tuer, scrupules, sans scrupules	Fataliste Optimiste incurable		Des principes (ne pas tuer de sang froid) Aucun remords	
SEXUALITÉ	Adorable avec les femmes, intéressé par une femme, malheureusement désir violent	Avant tout un homme		Pas certain de pouvoir résister au désir	Détachement et maîtrise totale de la fonction sexuelle
DIVERS	Enigmatique	Impénétrable	Etrange	Impénétrable	

celle qui porte sur chaque sujet séparément : on passe en revue les différents thèmes qu'il a abordés, ce qu'il a dit, et on tente, éventuellement, une synthèse individuelle. L'analyse horizontale, elle, traite chaque thème, relevant les différentes formes sous lesquelles il apparaît chez les personnes interrogées. Ces deux types d'analyses, simultanément, constituent bien un mode d'approche de notre problème, mais il reste tout à fait partiel, puisqu'on se contente de deux découpages particuliers du matériel, sans en obtenir une vue globale véritable.

La recherche de types idéaux tente de répondre à ce souci de globalité : en décrivant des types, avec leur cohérence interne, on rend chaque individu plus intelligible par sa relation aux différents types qui servent de points de référence. Ou, de façon moins prudente, on partage la population entre les différents types, en considérant les différences entre individus de la même catégorie comme négligeables ou non pertinentes. Malheureusement, malgré l'usage brillant, et convaincant, qu'a fait Max Weber de la méthode des types idéaux, il est resté très peu explicite sur la manière de les construire, et les applications ultérieures de la méthode relèvent plus de l'habileté du chercheur, de son savoir-faire, que d'une méthodologie communicable.

Pour progresser, et tenter d'y voir plus clair, il faut d'abord expliciter ce que pourrait être un discours satisfaisant portant sur un ensemble de textes, dans notre cas un ensemble d'entretiens. A notre sens, ce devrait être un discours prenant en considération à la fois les *traits communs* aux différents entretiens et leurs *différences*, en les organisant dans la mesure du possible.

Plaçons-nous dans le cas auquel peuvent se ramener de nombreuses analyses de contenu, où la méthode utilisée réduit chaque entretien à un ensemble de propositions. La démarche la plus élémentaire pourra consister dans la recherche de l'*intersection*, si elle existe, de ces ensembles, c'est-à-dire des propositions communes à tous les entretiens. On peut également élargir ce noyau commun en recherchant des points communs autres que les propositions, essentiellement des *relations* entre elles. Par exemple, dans les entretiens sur la représentation du psychologue, dont nous avons parlé plus haut, la plupart des personnes interrogées (pas toutes, mais cela n'est pas important pour l'exposé de la démarche) distinguent, et opposent, deux psychologies. Chez les uns ce sera l'opposition entre psychologie expérimentale et psychologie clinique ; pour d'autres entre psychologies scientifique et mi-scientifique ; d'autres encore distinguent une psychologie oppressive, au service de la société, et une psychologie libératrice, qui aide les individus, etc. Quel que soit le contenu des deux termes, les jugements de valeur qu'on porte sur chacun d'eux, on peut abstraire le trait commun constitué par l'opposition elle-même.

Même en élargissant ainsi la notion de trait commun, l'intersection des « contenus » de différents entretiens risque d'être faible. Or, il n'y a aucune raison d'admettre que seul ce qui est commun à tous mérite d'être retenu. C'est tout le problème du statut des différences inter (ou éventuellement intra) individuelles qui est posé [1]. Dans le cas des

1. Il est d'ailleurs rare qu'il soit posé explicitement. Soit on prend comme objet d'étude les différences inter-individuelles elles-mêmes, qui sont l'objet de la psychologie différentielle, soit on recherche des mécanismes généraux, et les différences apparaissent alors comme des obstacles à éliminer. La réduction d'une distribution à sa moyenne, l'emploi de méthodes statistiques et de modèles probabilistes constituent des réponses à ce problème.

entretiens, trois approches sont possibles. Premièrement, on peut éviter le problème en *postulant l'homogénéité* de l'ensemble des sujets, ou de certains sous-ensembles définis selon des critères extérieurs aux entretiens eux-mêmes. La deuxième approche consiste à *rechercher des types*, c'est-à-dire des sous-ensembles d'individus présentant plusieurs traits communs. Enfin, on pourra chercher à *organiser les différences*, c'est-à-dire considérer deux ou plusieurs traits différents comme des modalités distinctes d'une même variable.

Nous avons déjà fait allusion aux difficultés qui résultent de la non-distinction des discours d'individus différents. Néanmoins, il y a peut-être des cas où le postulat d'homogénéité, dûment explicité, peut se justifier, en particulier lorsqu'on a pris soin de constituer des sous-ensembles d'individus semblables selon des caractéristiques que l'on sait par ailleurs être importantes. Il faut bien voir que, lorsqu'on traite un ensemble de discours comme un discours unique, en faisant donc abstraction du locuteur, on admet que les différences entre eux sont aléatoires, et que c'est pour des raisons contingentes que chacun n'a formulé qu'une partie du discours commun.

La recherche des traits communs à des sous-ensembles de sujets, qu'on appelle ces sous-ensembles « types » ou qu'on préfère éviter les très nombreuses connotations de ce terme, élimine évidemment ces inconvénients. Au lieu de postuler l'homogénéité de groupes construits en fonction de critères antérieurs aux discours, c'est à l'intérieur de ceux-ci qu'on va rechercher des traits communs. C'est la méthode dite de la *recherche de concomitances*, dont nous avons déjà parlé. Mais, au lieu de nous intéresser aux traits eux-mêmes, et aux relations entre eux, nous nous servons des concomitances pour définir des groupes [1].

Pratiquement, la méthode des concomitances est lourde et exige, le plus souvent, le recours à l'informatique, surtout si le corpus est important. Ou alors, on ne recherche pas systématiquement toutes les concomitances, mais on se borne à vérifier des hypothèses préalablement posées à leur sujet, ce qui limite évidemment le nombre des opérations nécessaires.

Ces méthodes, par leur logique, rejoignent les techniques statistiques utilisées dans l'analyse des réponses à des questionnaires systématiques, bien que le problème de la légitimité de l'emploi d'inférence statistique se pose, et ne soit pas clairement résolu.

Toutefois, dans le cas des entretiens, et en particulier des formes les plus non directives, un problème spécifique apparaît : celui de la non-occurrence d'un trait. Dans le cas des questionnaires il arrive également que des personnes ne répondent pas à certaines questions ; mais on sait, dans ce cas, que la question leur a été posée, et, si l'enquête a été faite correctement, on sait en principe si cette absence de réponse est due à un refus, à l'ignorance, ou à d'autres raisons (cf.7.1.4.). Lorsque aucune question n'est

---

1. La *dualité individus-comportements* (ou, éventuellement, caractéristiques) se retrouve partout en psychologie, et dans toutes les méthodes sociologiques qui partent de données de base individuelles. Devant un matériel complexe qui associe à chaque sujet un ensemble de ses comportements, on peut toujours faire l'analyse soit en s'intéressant aux individus (par exemple, peut-on les ordonner d'un certain point de vue ?), soit en se centrant sur les comportements (par exemple, tel comportement accompagne-t-il toujours tel autre ?). Ce sont toujours deux manières d'analyser les mêmes données de base et une même structure. Mettre en évidence une relation entre individus, c'est montrer qu'une relation existe entre les ensembles de comportements de chacun d'eux. De façon duale, une relation entre comportement n'est rien d'autre qu'une relation entre les ensembles d'individus qui présentent ces comportements.

posée, l'absence d'occurrence d'un trait dans un entretien est beaucoup plus ambiguë ; en plus de toutes les raisons qui pourraient amener une personne à ne pas répondre à une question précise, il n'est pas possible d'exclure que le déroulement d'un entretien comporte des aspects contingents, qui peuvent orienter la chaîne associatrice vers certains thèmes plutôt que vers d'autres, ou des blocages inconscients, etc., sans qu'il nous soit possible, le plus souvent, de distinguer ces différents cas.

Il en résulte que la présence ou l'absence d'un trait ne peuvent pas être simplement considérés comme deux modalités d'une même variable. S'il est possible de définir un type par la co-occurrence de certains traits, le groupe complémentaire, formellement défini par leur « co-absence », risque d'apparaître comme moins satisfaisant, moins homogène, présentant une plus faible unité. En fait, ce sont des problèmes à résoudre dans chaque cas particulier, en fonction de ce que, par ailleurs, on sait, ou suppose, de l'ensemble du problème.

En général, lorsqu'on parle de « types », même sans référence à la méthode des types idéaux, on pense à des groupes qualitativement différents, à une simple classification des individus sans autre relation entre eux qu'une relation d'équivalence [1].

Mais on peut aussi introduire d'autres relations, soit entre les traits eux-mêmes, soit entre des types définis par plusieurs traits. C'est la troisième des approches que nous avons distinguées, celle qui cherche à mettre en évidence une structure des différences. Bien que d'autres puissent être envisagées, la plus simple et la plus couramment utilisée est simplement le pré-ordre total. Dire que des individus sont plus ou moins favorables à un certain objet, qu'ils accordent plus ou moins d'importance à tel autre, c'est dire qu'il est possible de les ordonner (en général pré-ordonner, puisqu'on admet, en principe, que deux individus puissent se situer au même niveau).

En général le matériel recueilli par entretien, quelle que soit la méthode d'analyse de contenu qu'on utilise, ne permet pas directement d'ordonner les sujets. Les discours sont trop hétérogènes, dans leur forme tout au moins, pour que des conclusions de type qualitatif, même s'il ne s'agit que du qualitatif faible de la relation d'ordre, puissent en être tirées. En revanche, il est possible, à partir d'entretiens, de faire des hypothèses sur l'existence d'une variable qualitative sous-jacente, une attitude par exemple, qu'on pourra ensuite chercher à opérationnaliser dans un questionnaire.

De même qu'un trait peut apparaître chez certains sujets et pas chez d'autres, une attitude peut être pertinente pour décrire les différences entre certains sujets, alors que pour d'autres tout ce qui est lié à cette dimension est absent. Nous retrouvons, à un niveau d'élaboration plus poussé, des problèmes que nous avons déjà rencontrés à propos des traits communs.

A la fin de la longue chaîne d'opérations diverses qui constituent l'analyse complète d'un ensemble d'entretiens, on dispose d'un texte nécessairement très éloigné des dis-

---

1. En admettant que les types constituent une véritable partition de l'ensemble des individus, c'est-à-dire que tel individu appartient à un groupe et à un seul. Dans les méthodes peu formalisées, dont il est question ici, ces conditions peuvent ne pas être remplies. Souvent, à cette phase de la recherche, on se contentera de définir clairement quelques groupes, en acceptant que le classement de certains sujets puisse être ambigu. Les types ainsi définis fonctionnent alors un peu comme des types idéaux. En revanche, si on essaie d'opérationnaliser les types par un questionnaire, il est indispensable d'arriver à une véritable partition.

cours initiaux. On a procédé à de nombreuses transformations, certaines exigeant des approximations diverses, d'autres recourant à des interprétations dont les fondements sont plus ou moins clairs. Des jugements de pertinence ont été portés, à chaque étape des distorsions ont pu se produire, des points importants ont pu être négligés. Afin d'éviter, au moins en partie, certaines de ces erreurs, et pour s'assurer qu'on a bien retenu et organisé tout ce qui peut sembler important, une dernière vérification est nécessaire. Elle consiste à retourner aux discours originaux, au texte même des entretiens, et à se demander, pour chacun, comment il se situe par rapport à la synthèse proposée.

C'est ainsi qu'on pourra s'apercevoir, par exemple, que certains traits, pourtant très importants au jugement immédiat, ont disparu au cours des diverses transformations ; que la synthèse finale comprend des affirmations qui ne sont fondées sur aucun entretien, mais résultent uniquement de l'influence des présupposés de l'analyste ; que certaines conclusions, qui dans la synthèse occupent une place considérable, ne proviennent que d'un seul entretien. Et ainsi de suite. Toutes ces erreurs, toutes ces distorsions, sont inévitables, quelle que soit la méthode utilisée, mais certaines techniques entraînent plutôt des erreurs par défaut (celles qui se veulent les plus objectives), alors que d'autres, au contraire, les plus interprétatives, ne fournissent que peu de garanties contre l'intrusion de présupposés extérieurs. Toutes risquent d'accorder un poids exagéré aux entretiens de certains sujets, aux dépens des autres ; que ce soient ceux qui sont les plus conformes aux hypothèses ou aux présupposés, ou simplement ceux qui parlent le mieux, disent le plus de choses, et les disent le plus clairement.

A ce propos, il faut souligner que le biais culturel est probablement un des plus graves qui risquent de fausser les conclusions tirées d'un ensemble d'entretiens libres. Toutes les formes d'enquête, par définition, recourent au langage, c'est entendu. Mais la technique des entretiens non directifs exige une production verbale importante. Or, on sait que la maîtrise du langage est très fortement liée au niveau culturel. Il y a donc des chances pour que les personnes les plus cultivées proposent des discours plus riches et plus clairs que les autres, ce qui, ensuite, risque de leur donner un plus grand poids dans les conclusions finales.

Bien qu'il ne s'agisse que d'un contrôle informel, le retour aux entretiens et leur confrontation avec la synthèse constitue donc un garde-fou nécessaire.

6.7. CONCLUSION

La conclusion de ce chapitre ne manque pas d'être évidente. Si, sur le plan pratique, les recherches produisent parfois des résultats intéressants pour une action limitée, au plan théorique les défaillances constatées laissent rêveur quant à l'outil et aux résultats qu'il peut produire.

Mais la voie étant tracée pour des recherches à venir, il n'en demeure pas moins que la pratique d'aujourd'hui existe et peut produire parfois quelques miracles.

BIBLIOGRAPHIE

- Chabrol C., Lacout A. et Verges P. et P., Palmade J., Roussen J., Laporte H., *« Interprétation et analyse de contenu II »*, Connexions, 1974, 12.

- Gardin J.-C., *Les analyses du discours*, Neuchâtel, Delachaux et Niestlé, 1974.

- Holsti O.R., « Content analysis », in Lindzey et Aronson E. (eds.) *Handbook of social psychology*, Reading, Mass., Addison Wesley, 1968. Vaste revue de questions sur les problèmes, méthodes et applications de l'analyse de contenu.

- Henry P., Moscovici S., « Problèmes de l'analyse de contenu », *Langages*, 2, Larousse, Paris, 1968. Interrogation théorique sur l'analyse de contenu.

- Pecheux M., *Analyse automatique du discours*, Paris, Dunod, 1969. Présentation d'une méthode d'analyse de corpus.

- Pecheux M., Fuchs C., Grevillon A., Henry P., Analyse du discours, langue et idéologie, *Langages 37* , Larousse, Paris 1975. Point sur l'analyse automatique du discours et mise en perspective théorique et épistémologique.

- Trognon A., Julien P. Levy A., Klein M. et Ramognino M., Loquet, Flahault F., *« Sujet (P) et Objet (P) de l'analyse de contenu (I) »*, Connexions, 1974, II.

Chapitre 7

L'ANALYSE STATISTIQUE DES QUESTIONNAIRES

7.1. LES TRAITEMENTS PRÉLIMINAIRES

7.1.1. La mise en forme des résultats

Le travail sur le terrain est terminé. Tous les questionnaires sont rentrés. Avant de procéder aux dépouillements proprement dits, c'est-à-dire aux opérations directement liées aux objectifs de l'enquête, un certain nombre de travaux préliminaires sont nécessaires, dont le but est de mettre les réponses à toutes les questions sous une forme homogène, de manière à pouvoir les traiter et vérifier l'existence de relations entre elles. Cela implique que les réponses à toutes les questions, quelle que soit leur forme ou leur contenu, se présentent de la même manière : à chaque question est associée une liste bien définie de réponses possibles, liste généralement assez brève. A chaque sujet on fera correspondre un, ou éventuellement plusieurs, éléments de cette liste, ses réponses ou des indicateurs qu'on aura construits à partir de ses réponses.

Ces premières opérations, sur lesquelles nous reviendrons, effectuées, quel que soit à ce moment l'aspect matériel des réponses recueillies (questionnaires remplis, feuilles de codage, cartes perforées, etc.) on peut les concevoir comme aboutissant à la formation d'un tableau à double entrée, avec les sujets en colonne et questions en ligne, ou l'inver-

se, la réponse d'un sujet à une question étant indiquée dans la case à l'intersection de la ligne et de la colonne correspondantes, (il ne faut pas confondre ce tableau global avec les tableaux croisés dont il sera longuement question plus loin).

Cette représentation, nous l'avons déjà rencontrée, mais moins explicitée, à propos du problème des analyses de contenu d'un ensemble d'entretiens (cf. 6.6.). Dans le cas des questionnaires, la situation est mieux définie, puisque c'est la construction même du questionnaire qui nous donne la signification des lignes, et non la diversité des entretiens individuels ; mais les problèmes principaux restent les mêmes.

sujets / questions	S_1	S_2	S_3	...	S_j	...	S_m
Q_1							
Q_2							
Q_3							
⋮							
Q_i					Rij		
⋮							
Q_n							

La similitude de présentation des différentes questions recouvre évidemment des contenus très différents. Ainsi les lignes du tableau pourront aussi bien être des caractéristiques individuelles telles que l'âge, le sexe ou le lieu de naissance, que des descriptions de comportements divers, ou des opinions, etc.

Si, formellement, toute enquête peut être représentée par un protocole de cette forme, qui regroupe la totalité de l'information recueillie, l'analyse porte très rarement sur ce protocole considéré dans son ensemble. Mises à part quelques techniques récentes comme les différentes formes d'analyse des données, dont nous parlerons rapidement plus loin, les dépouillements classiques, même ceux qui vont au-delà des pourcentages et des tableaux simples, consistent tous à analyser les relations entre un petit groupe de lignes : les hypothèses que l'on peut formuler font rarement intervenir un grand nom-

bre de variables. Chaque hypothèse portera donc sur les relations entre un petit nombre de lignes. Parfois, mais beaucoup plus rarement, les hypothèses peuvent porter sur l'existence de groupes de sujets (qu'on appellera éventuellement des types), et sur les relations entre ces groupes (par exemple lorsqu'on cherche à ordonner les sujets selon une attitude définie par les réponses à un grand nombre de questions). Dans ce cas l'analyse portera sur les relations entre colonnes, et non plus entre les lignes.

Dans la mesure où l'analyse ne porte pas sur le tableau dans sa totalité, on peut estimer qu'on perd de l'information puisque, lorsqu'on analyse les relations qui apparaissent entre, disons, Q1, Q2 et Q3, on néglige les relations qui peuvent aussi exister entre cet ensemble de variables et Qn. Toutefois, même si notre échantillon comportait suffisamment de sujets pour que l'analyse de telles relations complexes soit possible, nous rencontrerions de grandes difficultés à les interpréter.

L'analyse de données (analyse en composantes principales, analyse factorielle des correspondances, etc.) cherche précisément à saisir un ensemble de données complexes dans sa totalité, et à le faire sans poser d'hypothèses au départ. Mais si, par ces méthodes, il arrive effectivement qu'on obtienne une représentation très commode de l'ensemble des données recueillies, représentation qui peut être suggestive, elle ne dispense pas de la vérification d'hypothèses, peut-être plus limitées, mais qui pourront avoir une portée théorique et qui serviront de base à des généralisations (cf. 7.3.4.4.).

Actuellement, l'analyse des résultats d'une enquête se fait par ordinateur, ce qui permet de traiter rapidement et sans risque d'erreur des enquêtes complexes, portant sur un grand nombre de sujets. Les dépouillements les plus simples se présentent sous la forme de *tableaux de fréquences*, qu'il s'agisse simplement de la distribution des différentes réponses à une question (appelée souvent « tri simple » ou « tri à plat »), ou de *tableaux croisés*, donnant la fréquence des combinaisons de réponses à deux questions ou plus. On peut également obtenir des *listings* donnant la suite des réponses de chaque sujet à un grand nombre de questions, ou des calculs divers, allant de simples pourcentages à des moyennes ou à des tests statistiques. Enfin, il existe des programmes complexes permettant d'effectuer des analyses factorielles, de rechercher des échelles d'attitude ou des typologies, etc.

Certaines des conclusions auxquelles on veut aboutir (qu'il s'agisse d'estimations ou de vérifications d'hypothèses) s'expriment directement au moyen des réponses recueillies : on s'intéressera à la fréquence d'un certain comportement ou à la relation entre une caractéristique telle que, par exemple, l'âge, et ce même comportement. Dans d'autres cas, les hypothèses sont formulées en fonction de variables plus complexes, comme des attitudes, le niveau de vie, etc. Ces variables ne s'expriment pas directement par une seule réponse, mais on peut la construire à partir de plusieurs. Par exemple, on définira une attitude par la combinaison des réponses à plusieurs questions portant sur des opinions. Le mode de vie sera caractérisé par plusieurs comportements. Certaines hypothèses portent sur des typologies qui, par définition, sont construites à partir de plusieurs traits élémentaires. Dans tous ces cas la première étape du dépouillement devra consister à construire ces nouvelles variables, qui seront ensuite introduites dans les données de la même manière que les réponses obtenues directement des sujets.

En principe, le plan de dépouillement d'un questionnaire devrait se faire en fonction d'objectifs précis, qui auraient déjà guidé la conception et la rédaction du questionnaire. En fait, et même si ce n'est pas conforme à une image idéale de la démarche scientifique, une bonne partie des analyses effectuées sur les résultats d'une enquête correspondent plutôt au désir, encore vague, de « regarder ce qu'il y a », les hypothèses précises étant suggérées par l'examen et l'interprétation de tableaux simples.

Il est très rare que, dans l'analyse des résultats d'une enquête, on se borne à la constatation que telle hypothèse a été vérifiée et que telle autre ne l'a pas été. Il s'y ajoute toujours une phase d'*interprétation* des résultats obtenus. On ne se contentera pas, en général, de constater, par exemple, qu'un comportement varie avec l'âge, que ce soit ou non l'hypothèse qu'on avait faite, et a fortiori si on n'avait pas fait d'hypothèse, se contentant de se demander à quelles variables peut être lié le comportement en question. On se demandera le plus souvent pourquoi une telle relation existe, on se posera des questions sur les liens possibles entre une telle relation et d'autres qu'on aura éventuellement observées.

Cette phase d'interprétation, dans un processus qui se veut scientifique, pose problème. Tout d'abord par sa nature : en quoi consiste cette activité, sur quoi s'appuie-t-elle, quelle est sa validité ? Ensuite sur son rôle : pourquoi faut-il une phase distincte de la vérification des hypothèses ?

Nous avons vu que la conception même d'une enquête, et la formulation des hypothèses relevaient, plus que d'une déduction rigoureuse, de l'intuition ou de la créativité du chercheur [1]. Puis, parlant des différentes méthodes de questionnement, nous avons dit que nous ne disposions guère que d'une méthodologie négative, signalant des risques plus qu'indiquant ce qu'il faut faire. Enfin, l'interprétation, étape finale, nous apparaît aussi comme échappant à une démarche strictement scientifique. Seule subsiste la phase de vérification des hypothèses proprement dite ou d'analyse des données, où les techniques et les postulats sur lesquels elles s'appuient sont clairs.

L'analyse du processus même d'interprétation mériterait d'être étudiée non seulement en tant que méthode, dont on critiquerait la validité, mais d'abord en tant que comportement, en tant que pratique, dont il faut se demander en quoi ils consistent.

Dans ce qui suit, par la force des choses [2], nous parlerons plus, et de façon plus précise, de la vérification des hypothèses que de l'interprétation. Mais il est évident que les deux ne peuvent pas être dissociées.

Les techniques utilisées dans l'analyse des questionnaires ne sont rien d'autre que l'application, au cas particulier des données recueillies par enquête, des méthodes de la statistique descriptive ou inductive. Nous supposerons ces généralités connues pour l'essentiel, et ne reviendrons pas non plus sur la légitimité de leur emploi.

Pour pouvoir atteindre à l'homogénéité des données dont nous avons parlé au début de ce chapitre, les opérations suivantes peuvent être nécessaires :

— coder les questions ouvertes,

1. C'est à dessein que nous employons des termes vagues, à en être presque vides de sens, pour désigner des processus dont nous ne savons à peu près rien.
2. Qui comme chacun sait, est la forme que prend le destin chez les scientifiques.

— préciser comment on traitera les questions à choix multiples,
— préciser comment on traitera les différentes formes de non-réponse,
— construire des variables nouvelles à partir des données recueillies.

Ce dernier point est suffisamment important pour ne pas être inclus dans les préliminaires, bien qu'il en fasse logiquement partie. Nous lui consacrerons un chapitre entier.

Au début de la partie consacrée à la conception des questionnaires (cf. 4.1.) nous avons distingué quatre objectifs qui pouvaient être ceux de la phase quantitative d'une enquête :

— l'estimation de grandeurs absolues,
— l'estimation de grandeurs relatives (de proportions),
— la description d'une population ou de sous-populations,
— la vérification d'hypothèses.

Ces objectifs, s'il est utile de les distinguer, ne donnent pas une classification des techniques utilisées. On peut regrouper celles-ci en deux catégories qui interviennent dans la réalisation de ces objectifs :

— la construction de variables nouvelles (échelles d'attitude, typologies, indicateurs divers), et l'estimation des grandeurs correspondantes,
— l'analyse des relations et la vérification statistique d'hypothèses à leur propos.

C'est ce regroupement qui nous servira de plan pour l'exposé des méthodes.

7.1.2. La codification des questions ouvertes

Il s'agit, en fait, d'un problème d'analyse de contenu, avec la volonté d'en présenter les résultats sous une forme simple. En principe ce sont donc les mêmes techniques qui s'appliquent, ou au moins les mêmes orientations méthodologiques. Avec toutefois une différence : le corpus sur lequel on travaille, l'ensemble des réponses recueillies, est formé de discours en général très brefs. Le problème en est simplifié, mais des réponses brèves risquent, plus que des entretiens, d'être ambiguës, et ceci d'autant plus que, transcrites en général par l'enquêteur, et non enregistrées au magnétophone, elles risquent d'être incomplètes ou simplifiées.

Le travail de codification se fait en deux étapes : la première, par une analyse d'un échantillon du corpus, aboutit à l'élaboration du code lui-même. La seconde étape consiste à faire correspondre chaque réponse à une ou plusieurs catégories du code.

Pour être utilisable de façon satisfaisante, un code doit présenter plusieurs caractéristiques. Premièrement, le nombre des catégories qui le constituent ne doit pas être trop élevé, pour des raisons de commodité de manipulation, et surtout pour des raisons statistiques, qu'on retrouve d'ailleurs pour toutes les questions : lorsqu'on distingue trop de réponses, il reste trop peu de sujets dans chaque catégorie, et il n'est plus possible d'en tirer des conclusions significatives. Chaque fois que l'on est tenté de prendre en considération une différence, il faut se demander non si elle est intéressante en soi (la réponse a toutes les chances d'être affirmative), mais si elle est vraiment nécessaire pour les buts poursuivis. Sinon, en introduisant des catégories supplémentaires pour en tenir compte, on complique le travail des codeurs, et on masque l'information importante sous des nuances inutiles.

Deuxièmement, il faut s'assurer à la fois que toutes les catégories prévues sont effectivement utilisées, et que le nombre inévitable de réponses inclassables, reste très faible.

Troisièmement, les règles définissant l'attribution d'une réponse à une catégorie doivent être aussi explicites que possible, et laisser le minimum de champ au jugement du codeur. Dans la suite du travail, lorsque les questions ainsi codées seront insérées dans des analyses plus complexes, il faudra conserver à l'esprit ces règles, qui constituent la véritable définition des catégories.

Quatrièmement, il faut préciser si chaque réponse doit être attribuée à une seule catégorie, ou si l'on admet la possibilité de coder une même réponse dans plusieurs catégories. Mais se donner cette possibilité n'est utile que si l'on est amené à l'utiliser souvent. Si les réponses multiples restent l'exception, il est préférable d'accepter de perdre une certaine information, et faire un choix parmi les diverses catégories envisagées. Nous reviendrons sur ce point au chapitre suivant.

Le code doit être élaboré par le chercheur lui-même ; il est seul à pouvoir distinguer, dans la multiplicité et la diversité des réponses, ce qui est nécessaire pour les analyses qu'il a l'intention d'effectuer. En revanche, le codage lui-même doit être effectué par d'autres, pour éviter que le poids des hypothèses ne vienne biaiser les décisions, des jugements subjectifs devant des cas non explicitement prévus étant rarement évitables.

Lorsqu'on en a les moyens, faire coder un même questionnaire indépendamment par deux codeurs, ou même plus, apporte une sécurité supplémentaire. C'est d'autant plus nécessaire que, l'expérience le montre, la concordance entre codeurs est souvent médiocre. Le double codage permet de repérer les cas où il y a désaccord, de s'interroger sur les causes, et de prendre les mesures nécessaires, qui pourront comprendre une amélioration du code lui-même.

Chaque question doit être codée indépendamment des autres. Si, par exemple, pour simplifier le travail, et aussi pour rendre la tâche des codeurs moins monotone, on fait coder toutes les questions ouvertes d'un même questionnaire les unes à la suite des autres, il y a de grands risques pour que le codeur fasse appel, dans les cas douteux, à sa conception de la cohérence des réponses, ou, sans s'en rendre compte, devienne victime d'un effet de halo. Les relations qu'on pourra ensuite constater entre les réponses ne feront que traduire cette cohérence introduite par les codeurs.

A côté des questions ouvertes proprement dites, qui exigent donc un codage souvent délicat et complexe, la préparation des réponses comprend souvent un codage plus simple pour passer des réponses données aux catégories souhaitées pour l'analyse.

Par exemple, on demande en général aux personnes interrogées d'indiquer leur profession de façon précise et détaillée. Mais ce qu'on veut en définitive, ce sont les Catégories Socio-Professionnelles (C.S.P.) de l'INSEE. Les détails sont demandés uniquement pour éviter les erreurs ou les ambiguïtés, bien des personnes indiquant spontanément leur profession en des termes qui ne permettent pas le passage aux C.S.P. Ce codage devra donc être effectué plus tard.

De même, on demandera souvent l'âge à un an près, même si ensuite on se contente de catégories moins fines, à cinq ou même dix ans près.

De façon générale, on procède souvent de cette manière pour toutes les informations quantitatives, ce qui permet, s'il y a lieu, de recourir à différents regroupements.

Demander des informations précises peut être utile lorsqu'on prévoit plusieurs codages. Par exemple, on pourra demander le lieu de naissance exact, puis coder d'une part la région, d'autre part la taille de l'agglomération. Demander simplement le nom de la commune de naissance, puis faire le double codage, peut être plus sûr que de poser directement les deux questions.

Enfin, certaines variables nouvelles peuvent être contruites par combinaison de réponses à plusieurs questions. Par exemple, on peut vouloir utiliser la variable « cycle de vie », obtenue par combinaison de l'âge, de l'état matrimonial, du nombre et de l'âge des enfants. Contrairement au codage des questions ouvertes, qui exige presque toujours une part de jugement et d'interprétation, ces combinaisons peuvent se faire soit à la main, soit automatiquement, tous les programmes de dépouillement d'enquêtes sur ordinateur comportant la possibilité de créer des variables nouvelles à partir des réponses obtenues.

Dans un cas comme dans l'autre il est indispensable de donner aux codeurs, ou au programmeur, la définition précise de chaque catégorie de la variable nouvelle, mais aussi des indications précises concernant ce qu'il faut faire dans les cas exceptionnels. Dans l'exemple du cycle de vie, que fera-t-on des célibataires d'âge moyen ? Des personnes seules avec enfant ? En principe, il faut passer en revue toutes les combinaisons concevables, même les moins vraisemblables, et décider ce qui devra en être fait, soit qu'on les mette dans une rubrique « inclassables », soit qu'on les attribue à une catégorie pas trop éloignée.

Dès que nous sortons des quelques cas des variables nouvelles simples, comme le style de vie, nous abordons un domaine très vaste, celui de la construction d'indicateurs divers à partir des réponses. Ce n'est plus un simple problème d'instructions de codage, mais la conception même de ces variables et de leurs indicateurs pose des problèmes à la fois fondamentaux et de technique, dont nous parlerons en détail plus loin.

7.1.3. Les questions à réponses multiples

Lorsqu'on pose une question fermée avec plus de deux réponses possibles, il arrive qu'on ne veuille pas imposer au sujet de se limiter à choisir une seule parmi les réponses qui lui sont proposées. Nous avons vu que plusieurs consignes différentes sont possibles. Rappelons-les :

— choisir un nombre libre de réponses,
— choisir un nombre imposé de réponses,
— choisir et ordonner un nombre imposé de réponses,
— ordonner la totalité des réponses proposées.

Lorsque la question est ouverte, il est possible qu'on appauvrisse trop les réponses en les codant une seule fois ; ou alors on est amené à construire un code ayant un trop grand nombre de rubriques, et qui en devient vite inutilisable, sauf évidemment lorsqu'on dispose d'un très grand nombre de sujets. Il est préférable de scinder les réponses obtenues et de considérer qu'une même personne a donné plusieurs réponses. Ce faisant, on perd évidemment l'information fournie par le fait que c'est un même sujet qui a énoncé les différents thèmes. Comme souvent, il y a un arbitrage à faire entre la richesse de l'information retenue et les possibilités de traitement de celle-ci.

Quelle que soit la forme initiale de la question, ouverte ou fermée, on se trouve donc finalement associer à chaque sujet un ensemble de réponses ordonnées ou non.

Une première manière de traiter ces réponses consiste à rester très près des données et à associer effectivement à chaque sujet sa combinaison de réponses. On aurait donc une application de l'ensemble des sujets, selon le cas, dans l'ensemble des parties de l'ensemble des réponses, ou dans l'ensemble des combinaisons des réponses n à n (si n est le nombre de réponses imposé), ou encore dans l'ensemble des permutations des n réponses, si c'est un ordre qui était demandé.

Cette manière de faire présente des avantages appréciables, et un inconvénient très grave. Premièrement, on respecte ainsi au maximum ce que le sujet a effectivement dit : s'il a nuancé une des réponses en en choisissant une autre en même temps, on en tient compte. Deuxièmement, on fait correspondre à chaque sujet un seul élément du nouveau code ainsi constitué, ce qui facilite beaucoup les manipulations ultérieures. En revanche, et c'est là un inconvénient majeur, on émiette la population en un très grand nombre de catégories, chacune contenant donc très peu de sujets, et il devient presque impossible d'en tirer des conclusions quelconques, sauf s'il se trouve que les réponses sont très regroupées. Si, par exemple, on a proposé cinq réponses, ce qui n'est pas beaucoup, avec choix libre il y a $2^5 = 32$ combinaisons de réponses possibles (31 s'il est exclu qu'un sujet n'ait choisi aucune réponse). Répartir les sujets entre 31 catégories, surtout si on les croise ensuite avec d'autres variables, est le plus souvent impossible, étant donné la taille habituelle des échantillons.

C'est pour cela que, lorsqu'on demande simplement au sujet des choix, et non un ordre de préférence, on traite habituellement chaque réponse proposée comme une question indépendante des autres, auxquelles le sujet aurait répondu par oui ou par non. Ce qu'on perd en information sur les réponses simultanées, on le gagne en possibilités d'utilisation.

On peut d'ailleurs récupérer une partie de l'information perdue par des codages complémentaires, par exemple en codant le nombre de réponses données par chaque sujet, lorsque celui-ci n'est pas imposé. Dans les questions d'opinion, cette valeur est peut-être peu significative, mais lorsqu'il s'agit de connaissances, par exemple lorsqu'on demande à un sujet de désigner, dans une liste de produits, ceux qu'il consomme, le nombre de produits mentionnés peut être utile à connaître. En croisant cette nouvelle variable avec la question initiale, on pourra en outre savoir, pour chaque produit, s'il est consommé par des personnes qui en mentionnent beaucoup ou qui en mentionnent peu, ce qui est une des informations qu'on aurait pu tirer de l'analyse des réponses complètes.

Lorsqu'on traite ainsi chaque réponse séparément, on peut se demander s'il faut calculer le pourcentage de choix de chaque réponse par rapport au nombre total de sujets ou au nombre total de réponses. C'est en général la première solution qui est adoptée : pour chacune des réponses possibles, on indique le pourcentage de sujets qui l'on choisie. Les cas où c'est la seconde manière de faire qui serait la plus adaptée sont très rares.

Les résultats devraient donc se présenter sous la forme du tableau suivant, où N représente le nombre de sujets interrogés.

	CHOISIE		NON CHOISIE		TOTAL
Réponse 1	f_1	p_1 %	$N-f_1$	100	N = 100 %
Réponse 2	f_2	p_2 %	$N-f_2$	$100-p_2$ %	N = 100 %
Réponse 3	f_3	p_3 %	$N-f_3$	$100-p_3$ %	N = 100 %

Pour des raisons de commodité, et d'homogénéité de présentation avec les questions à choix unique, les tableaux publiés sont présentés sous une autre forme :

Réponse 1	f_1	p_1 %
Réponse 2	f_2	p_2 %
Réponse 3	f_3	p_3 %
⋮		
Total	N	100 %

ce qui exige de préciser que la somme des pourcentages dépasse cent. Sans être fausse, une telle présentation peut être trompeuse, car elle masque le fait que, en principe, chaque réponse a été choisie ou rejetée indépendamment des autres.

Il arrive qu'on impose le nombre de réponses que le sujet pourra choisir ; c'est généralement dans l'intention de rendre les résultats de tous les sujets plus homogènes, d'éviter que certains sujets s'arrêtent à la première réponse qui leur paraît à peu près satisfaisante, ou que d'autres les choisissent toutes pour éviter de se prononcer. Mais cette consigne est artificielle, et de plus on ne sait pas traiter cette situation autrement que lorsque le nombre de réponses est libre ; en d'autres termes, on n'a pas de moyen de tenir compte formellement, dans l'analyse des résultats, de cette contrainte supplémentaire. Sauf cas particulier cette consigne nous semble donc peu intéressante.

Demander aux sujets d'ordonner les réponses proposées en fonction de leurs préférences, par exemple, apporte également une information riche, mais peu utili-

sable. Prenons d'abord le cas, le plus simple, où on demande d'ordonner toutes les réponses. A chaque sujet correspond donc une permutation.

On peut faire les mêmes remarques que précédemment concernant le nombre de catégories ainsi reconstituées. Avec 5 réponses proposées, il y a 5 ! = 120 formulations possibles, donc 120 catégories, ce qui est évidemment beaucoup trop.

Si nous nous intéressons à la structure des préférences, c'est-à-dire à la manière dont celles-ci s'organisent, aux critères qui peuvent en rendre compte, à l'homogénéité des sujets de ce point de vue, on peut s'appuyer sur toutes les méthodes, algébriques ou vectorielles, développées au cours des dernières années. Nous n'en parlerons pas ici, l'exposé de ces méthodes exigeant des développements qui seraient disproportionnés avec leur importance dans le cadre des techniques d'enquête proprement dites.

Parmi les méthodes simples, on peut mentionner l'attribution à chaque réponse de son rang moyen. Mais cela revient à traiter les rangs comme une variable numérique et à admettre que, pour tous les sujets, l'intervalle entre deux rangs successifs est toujours le même. Enfin, l'ordre ainsi obtenu sur les réponses, celui des moyennes des rangs qui leur sont attribués, peut n'avoir guère de sens : résultat peut-être d'une compensation entre des ordres très différents, il se peut que personne n'ait manifesté cet ordre de préférence « moyen ». On aurait tort, sauf si on s'est assuré au préalable que les ordres individuels sont suffisamment proches les uns des autres, par exemple au moyen du coefficient de concordance de Kendall, de penser qu'un tel ordre moyen représente d'une manière quelconque les préférences du groupe des sujets interrogés.

On peut également, ce qui est un peu plus riche et risque moins d'induire en erreur, indiquer pour chaque réponse le nombre de sujets qui l'on choisie en premier lieu, en deuxième lieu, etc. On obtient alors le tableau suivant :

	Rang 1	Rang 2	Rang 3.....	
Réponse 1	f_{11} p_{11}	f_{12} p_{12}		N = 100 %
Réponse 2	f_{21} p_{21}	f_{22} p_{22}		N = 100 %
Réponse 3	f_{31} p_{31}	f_{32} p_{32}		N = 100 %
⋮	N = 100 %			

Le tableau peut se lire en ligne : sur 100 sujets, p11 ont attribué à la réponse 1 le rang 1, p12 le rang 2, etc. En revanche, la lecture en colonne est plus difficile. La première colonne ne pose pas de problème, et peut se lire : sur 100 sujets, p11 ont attribué le rang

1 à la réponse 1, p21 à la réponse 2, etc. Les colonnes suivantes sont plus délicates. Le nombre de sujets ayant attribué le rang 2 à la réponse 1 doit être rapporté non au nombre total de sujets, mais à ceux qui pouvaient encore le faire, c'est-à-dire à ceux qui ne lui avaient pas attribué le rang 1, donc n-f11. On considérera donc un pourcentage p'12 = f12N-f11, qu'on lira : sur 100 personnes n'ayant pas attribué le rang 1 à la réponse 1, p'12 lui ont attribué le rang 2.

On voit que l'analyse devient rapidement complexe, et les résultats difficilement interprétables. C'est pourquoi on voit souvent des résultats d'enquêtes qui se contentent de la première colonne, c'est-à-dire des attributions du rang 1. Autrement dit on se ramène au cas où l'on impose à chaque sujet de ne donner qu'une seule réponse. Il était alors inutile de lui en demander plus.

Si nous nous sommes arrêtés un peu longuement sur des procédures que nous critiquons, et en définitive déconseillons, c'est que leur emploi est assez fréquent, et surtout c'est qu'il est représentatif d'une attitude très courante, qui consiste à s'efforcer d'obtenir le maximum d'informations, sans prévoir de façon précise comment elles seront utilisées. La démarche est banale : on élabore une liste de réponses, on se demande s'il n'est pas artificiel de limiter les possibilités d'expression des sujets en leur imposant de ne choisir qu'une seule réponse, et par conséquent on élargit la consigne, l'ordre étant censé apporter encore des nuances supplémentaires. Puis, devant les difficultés de traitement, on appauvrit l'information recueillie, se ramenant à ce qui avait été initialement envisagé, puis rejeté. Il aurait été préférable de commencer par-là.

Cela ne veut pas dire qu'il faille condamner toutes les consignes complexes : il faut simplement, et ce devrait être une règle absolument générale, toujours se demander, avant de poser une question et de décider de sa forme, ce qu'on en fera, et à quoi les réponses recueillies serviront.

Lorsqu'on a laissé les sujets libres du nombre de réponses qu'ils peuvent choisir, on peut, nous l'avons vu, traiter chaque réponse comme une question indépendante, à laquelle on aurait répondu par oui ou par non. Vues sous cet angle, les questions à réponses multiples ne posent aucun problème spécifique, puisqu'elles se ramènent à autant de questions binaires qu'on avait proposé de réponses. Mais l'analogie est formelle. En effet, lorsqu'on présente simultanément à une personne plusieurs réponses, elle les compare, et ses choix se font par référence aux autres. Si les questions avaient été réellement posées sous forme indépendante, c'est à chacune, en principe, qu'elle aurait répondu, le poids des autres étant beaucoup plus faible. Par exemple, on peut très facilement imaginer un sujet qui à chaque question indépendante réponde « non », alors que, si les mêmes énoncés lui étaient présentés comme des réponses entre lesquelles choisir, il en aurait très certainement désigné au moins une, pour éviter ce qui serait apparu comme une non-réponse. Cela n'implique pas, formellement, de traiter les deux situations de manière différente, mais il faudra garder la situation à l'esprit au cours des interprétations. De plus, si l'on traite chaque réponse isolément, on perd l'information apportée par les structures de réponses. Il faut apprécier leur importance ou leur nécessité.

7.1.4. Les non-réponses

Que la question posée soit ouverte ou fermée, il est rare que toutes les réponses recueillies entrent dans les catégories prévues, et puissent être décomptées avec les autres. Dans la présentation des résultats, on les regroupe le plus souvent en une seule catégorie résiduelle, éventuellement dans deux, sous les rubriques « sans opinion » ou « ne sait pas », et « autres réponses ».

Il est important de ne pas traiter ces « non-réponses » comme un simple résidu sans intérêt. Leur proportion peut être très significative ; par exemple, on sait qu'elle peut varier considérablement en fonction de la catégorie sociale, de la position par rapport au thème traité, etc. Ne raisonner que sur la sous-population qui entre dans la grille prévue amène à généraliser abusivement ses réponses à l'ensemble de la population. D'autre part, ces « non-réponses » peuvent être très diverses, et il serait souhaitable, pour une interprétation correcte des résultats, de faire un certain nombre de distinctions parmi elles :

— Les réponses « je ne sais pas », ou « je n'ai pas d'opinion », explicitement formulées.
— Les refus de répondre, parce qu'on juge la question indiscrète.
— Les refus de répondre, parce qu'on juge la question mal posée, ou tendancieuse.
— Les questions sautées ou oubliées par l'enquêteur.
— Les refus de répondre, en fin de questionnaire, parce qu'on n'a plus le temps.

Lorsque la question est ouverte il faut ajouter :

— Les réponses qui indiquent manifestement une incompréhension de la question.
— Les réponses qu'on n'a pu coder dans aucune des catégories prévues.

Et pour les questions fermées :

— Les réponses autres que celles qui sont proposées.
— Le refus de toutes les réponses proposées, mais sans que le sujet en donne une autre.

Il est évident que les significations de ces types de « non-réponse » sont très différentes et que, en les confondant, on perd de l'information. Il est toutefois souvent difficile de les distinguer, parce qu'il est rare qu'on dispose de toute l'information voulue, qui serait la transcription textuelle de tout ce qu'a dit le sujet quand on lui a posé la question. Or, dans un questionnaire statistique, les enquêteurs ont tendance à travailler rapidement, et à ne pas noter les détails dont, faute d'information suffisante, ils ignorent l'utilité. Si l'on veut pouvoir faire ces distinctions, il faut donner aux enquêteurs des consignes très strictes dans ce sens.

On est aussi souvent amené à confondre les différents types de « non-réponse » parce que chaque catégorie est peu représentée, et que les quelques personnes qu'on pourrait y placer seraient, de ce fait, inanalysables. Selon le cas, d'ailleurs, l'importance de chacune de ces catégories pour l'interprétation des résultats est très variable.

Pour certains problèmes, les « sans opinion » ou ceux qui disent « je ne sais pas », sont intéressants en eux-mêmes. C'est une réponse comme une autre. C'est évidemment le cas des enquêtes d'opinion, où ce type de réponse peut être très important, et varier

de façon significative ; au cours d'une étude de marché, le fait qu'une personne ignore quelle marque d'un produit elle utilise, ou son prix, est également utile.

Pour d'autres cas, une proportion notable de « ne sait pas » indique plutôt que la question a été mal posée, ou du moins qu'il y a intérêt à la modifier. C'est le cas, par exemple, des questions posées en vue de la construction d'une échelle d'attitude. Dans la mesure où ces questions ne sont pas intéressantes en elles-mêmes, mais seulement parce qu'elles permettent, combinées avec d'autres questions, de classer les sujets, on n'aura aucun intérêt à les conserver si le taux de non-réponses est important.

Enfin, lorsqu'une question est posée pour obtenir, par l'intermédiaire des sujets, une information objective, et non la représentation qu'on s'en fait, constater que certaines personnes sont incapables de répondre n'apporte aucun renseignement intéressant. C'est le cas, par exemple, des enquêtes économiques ou sociologiques sur les budgets-dépenses ou les budgets-temps, la fréquentation d'équipements, les déplacements, l'activité professionnelle, le comportement électoral, etc. Le fait que certaines personnes se disent incapables de répondre à une question de ce type n'a aucun intérêt dans la perpective selon laquelle sont conçues ces enquêtes, et diminue la qualité des résultats. On peut essayer, dans de tels cas, de recourir à des aides diverses pour améliorer l'évocation et recueillir une meilleure information, de faire noter les comportements pendant une certaine période, etc., de manière à réduire autant que possible aussi bien les oublis que les autres formes de distorsion.

Malheureusement pour l'interprétation, la réponse « je ne sais pas » ne correspond pas toujours à une absence d'opinion ou à une ignorance réelles. Certaines réactions qui correspondent plutôt à l'une des autres catégories de non-réponses s'expriment facilement de cette manière, soit parce que c'est la réponse la plus facile pour le sujet, soit parce que, dans le doute, c'est dans cette catégorie apparemment la plus neutre, que les enquêteurs préfèrent coder une réponse vague ou équivoque.

La réponse « je ne sais pas » peut ainsi camoufler un refus conscient de répondre, lorsqu'on n'ose pas le faire ouvertement. Mais elle peut aussi traduire l'impression plus ou moins vague que la question est mal posée ou qu'elle est trop compliquée, ou encore qu'il faudrait réfléchir trop longtemps pour répondre. On ne peut pas s'attendre à ce que les personnes interrogées fassent spontanément les distinctions que nous sommes en train de développer ici, sauf peut-être si on les leur propose explicitement.

Les refus de répondre à une question sont relativement rares, dans les questionnaires courants. Celui qui a accepté de se prêter à l'enquête répond en général à toutes les questions, ou, s'il ne veut pas répondre, préfère dire qu'il ne sait pas. Il y a toutefois une question à laquelle le taux de refus est habituellement très élevé (de l'ordre de 20 % dans l'ensemble de la population, plus dans certaines catégories) : c'est celle qui porte sur les revenus, qui est souvent jugée très indiscrète, et qu'on place en général, pour cela, à la fin du questionnaire, pour ne pas en perturber la passation.

Les refus explicites d'une question parce qu'on la juge mal posée, ou parce que les réponses proposées ne sont pas satisfaisantes, sont également rares. Un faible taux de refus n'est toutefois pas une garantie suffisante de validité de la question. En effet, il est rare qu'une question soit posée de façon tout à fait absurde. Sans être conforme à la manière dont certaines personnes posent le problème, la formulation est le plus souvent

suffisamment acceptable pour que les sujets y répondent. De plus, expliquer en quoi la question est mal posée est évidemment difficile et ne correspond pas à l'attitude relativement passive que la plupart des questionnaires induisent chez les sujets.

On pourrait évidemment essayer d'atténuer ce conformisme en incluant dans la liste des réponses proposées, des catégories correspondant aux différentes formes de refus, pour que les personnes interrogées sachent que c'est permis, que cela fait aussi partie de la règle du jeu. Mais on court alors le risque de voir une proportion trop forte de sujets se réfugier dans cette forme d'abstention, plutôt que de réfléchir à la meilleure réponse.

La distinction entre les « ne sait pas » et les refus est particulièrement importante dans les enquêtes politiques, surtout les sondages d'opinion. On propose aux sujets, en général, le choix entre un petit nombre de solutions à un problème, sans leur dire si on leur laisse ou non la possibilité de refuser ces alternatives et de suggérer des solutions tout à fait différentes. Or, s'ils le font, ils seront, lors de la publication, confondus avec les « ne sait pas ».

Lorsque la question est posée sous forme ouverte, c'est au moment du codage qu'il faut introduire ces distinctions, et en plus prévoir une catégorie supplémentaire, les « autres réponses », celles qu'il est impossible de classer dans aucune des catégories prévues, et qui sont chacune trop peu fréquentes pour qu'il vaille la peine de constituer de nouvelles catégories pour elles. Le groupe « autres » est donc évidemment très hétérogène, mais s'il est peu nombreux, ce n'est pas grave. Si toutefois on est obligé d'y classer beaucoup de réponses, il faut se demander si on n'a pas fait des distinctions trop fines, et s'il n'est pas possible, malgré tout, de la subdiviser en un petit nombre de catégories plus homogènes. Il faut aussi distinguer ces réponses inclassables, mais de même type que les autres, de celles qui se situent à un autre niveau, comme les refus ou les critiques de la question.

La liste des réponses proposées à une question fermée comprend, en général, également la rubrique « autres réponses », l'enquêteur ayant pour consigne de faire expliciter cette réponse et de la transcrire textuellement. Parfois, elle ne diffère que par des détails d'une des réponses proposées ; dans ce cas, on la confondra évidemment avec celle-ci. Si ce n'est pas possible, il faut en faire l'analyse systématique, et décider ce qu'on en fera à partir de cette analyse, si le cas se présente suffisamment souvent.

Toute présentation des résultats d'enquêtes devrait donc distinguer dans la mesure du possible, les différentes formes de non-réponses, ce qui implique que les enquêteurs enregistrent les informations adéquates. Ne pas le faire peut introduire des erreurs graves. Dans les sondages d'opinion, par exemple, il est grave de confondre ceux qui disent n'avoir aucune opinion et ceux qui, critiquant la perspective politique sous-jacente à la formulation de la question, refusent d'y répondre.

Cela devrait retentir également sur la manière de calculer les pourcentages de réponses. Si les non-réponses ont un sens, il faut évidemment prendre comme base de calcul la totalité de l'échantillon. Calculer le pourcentage sur la seule partie qui a donné une des réponses prévues n'a guère de sens, puisqu'on ne sait pas à quelle population généraliser le résultat. Si on veut l'étendre à l'ensemble de la population initiale, cela revient à supposer que, parmi ceux qui n'ont pas répondu, les différentes réponses se

répartissent de la même façon que parmi ceux qui ont répondu, ce qui est un postulat parfaitement injustifié.

Lorsqu'on recherche une information objective, la situation est différente. Les calculs n'ont de sens que par rapport à ceux qui ont effectivement donné une réponse. Mais, dans ce cas, on sait que l'échantillon est biaisé, puisque rien ne nous assure que ceux qui n'ont pas répondu ont eu le même comportement que ceux qui ont accepté de répondre. Si l'on veut déterminer, par exemple, le taux de lecture d'un journal par la question « avez-vous lu le dernier numéro de... ? », il est évident qu'on ne l'obtiendra pas en calculant le pourcentage de « oui » par rapport à la totalité de l'échantillon. Ce serait, dans ce cas, admettre que ceux qui n'ont pas répondu ne l'ont pas lu. On doit donc se contenter du pourcentage par rapport au nombre de réponses, mais en sachant que le sous-échantillon utilisé (celui des personnes ayant répondu « oui » ou « non ») est certainement biaisé par rapport à la population totale. Il est en effet probable que ceux qui lisent le journal en question tout à fait régulièrement, ou jamais, répondent, ainsi que ceux qui, bien que ne le lisant qu'occasionnellement, y attachent de l'importance. La fréquence estimée ne sera donc vraisemblablement pas correcte.

Analyser les non-réponses, voir comment elles se répartissent dans différentes catégories, comment leur fréquence est liée à d'autres informations recueillies dans le questionnaire peut présenter un intérêt en soi. Répondre ou ne pas répondre à une question peut être, comme toute autre réponse, un indicateur de l'attitude à l'égard du problème abordé, mais aussi, ce qui complique l'interprétation, de l'attitude à l'égard de l'enquête et de la relation du sujet avec l'enquêteur.

7.2. LES VARIABLES

7.2.1. La construction des variables

7.2.1.1. Observables, indicateurs et variables intermédiaires

Parmi les réponses obtenues à un questionnaire, certaines peuvent être directement pertinentes pour le problème qu'on cherche à traiter, alors que d'autres ne nous intéressent que par leurs relations, supposées ou postulées, avec d'autres variables, qu'il n'est pas possible de saisir directement.

Lorsqu'on interroge quelqu'un sur son âge, sa profession, ses dépenses de la veille, ou sur son opinion sur la peine de mort, ce sont bien ces variables elles-mêmes qu'on veut connaître : ce sont elles dont on veut estimer la distribution, ou qui interviennent dans la formulation des hypothèses qu'on veut vérifier. Dans d'autres cas, lorsqu'on veut étudier une attitude, par exemple, un style de vie, ou un niveau de satisfaction, ces variables, qui sont celles qui nous importent, ne sont, par définition, pas directement accessibles.

Cette distinction ne recouvre pas celle que nous avons faite entre les questions de fait et les questions d'opinion. Les unes comme les autres peuvent nous servir à saisir indirectement des variables sous-jacentes. Par exemple, selon le problème traité, on cons-

truira une attitude à partir de questions d'opinion, d'une énumération de comportements, ou d'une combinaison des deux. Une variable comme la sociabilité ne se mesure en général que par des questions de fait, le nombre d'amis ou de contacts, par exemple.

Une attitude, qu'on la définisse comme un état intérieur ou comme une propension à agir d'une certaine manière, n'est pas elle-même observable. Selon l'obédience épistémologique ou méthodologique à laquelle on se rattache, on dira que c'est une *variable sous-jacente*, *intermédiaire* ou *latente*, une *construction théorique* ou *hypothétique*, ou encore un *concept théorique*. Tous ces termes, en fait équivalents, désignent des notions qu'on ne peut pas saisir directement, mais seulement *inférer* ou *construire* (du point de vue épistémologique le choix de l'un ou l'autre de ces termes est significatif) à partir des *données manifestes* ou des *observables* qui sont, dans le cas particulier des enquêtes, les réponses obtenues. Celles-ci ne présentent donc pas d'intérêt intrinsèque, mais ne sont utiles qu'en tant que symptômes ou indicateurs de variables intermédiaires. Les hypothèses ne sont, le plus souvent, pas directement exprimées en termes de réponses, mais au moyen de ces variables intermédiaires. La liaison entre celles-ci et les observables constitue donc un problème fondamental.

Dans le premier des cas que nous avons distingués, lorsque les réponses sont immédiatement pertinentes, elles peuvent être acceptées telles qu'elles ont été obtenues, les seuls risques d'erreur ou de biais pouvant provenir d'une mauvaise compréhension de la question par le sujet, de son ignorance, d'oublis, de déformations liées à la situation d'enquête. Toutes les précautions que nous avons énumérées dans la partie consacrée à la construction des questionnaires visent à réduire ces différentes erreurs. Dans le second cas, ces problèmes existent évidemment aussi mais il faut en outre se poser des questions sur la validité de la correspondance admise entre les réponses et la variable intermédiaire dont on a supposé l'existence.

Comme beaucoup de distinctions, celles que nous faisons entre observables et variables intermédiaires d'une part, entre réponses directement utiles et indicateurs d'autre part, si elles sont nettes dans leur principe, ne le sont pas toujours dans les cas particuliers. Lorsque les hypothèses sont clairement formulées, et déduites d'une théorie explicite, distinguer sans ambiguïté ce qui appartient à un niveau ou à l'autre est possible. Mais, dans la pratique, on se trouve rarement dans une situation aussi favorable, et il n'est pas toujours facile de savoir si une réponse nous intéresse en elle-même ou par sa relation avec une autre variable. Par exemple, des variables comme l'âge ou la catégorie socio-professionnelle ne sont souvent utilisées, en fait, que comme des indicateurs d'autres facteurs. Par exemple, la catégorie socio-professionnelle représentera un certain niveau culturel, une tranche de revenu, un style de vie, des habitudes de consommation et de loisirs, etc. Selon le problème, l'âge sera interprété comme un état physiologique, des rôles sociaux différents, des étapes de carrière, etc., ou beaucoup d'autres composantes, souvent beaucoup plus difficiles à expliciter (cf. 7.2.2.2.). Mais si ces facteurs « latents » interviennent de façon non formalisée dans les interprétations qui seront proposées, les relations obtenues, les hypothèses elles-mêmes font le plus souvent intervenir les variables « objectives » en question, et non ce qu'elles représentent.

De même, une question portant sur une opinion pourra être posée pour elle-même, par exemple lorsqu'on veut connaître le nombre des Français favorables à la peine de

mort. Mais, dans des analyses plus complexes, le chercheur, devant la nécessité d'interpréter les corrélations où cette même opinion intervient, pourra s'appuyer sur les connotations qu'évoque pour lui cette réponse, et, par exemple, pourra considérer que ceux qui se disent favorables à la peine de mort ont des chances d'être plus conservateurs, ou plus autoritaires, que ceux qui s'y déclarent hostiles.

Pour poser le problème en termes suffisamment généraux, et pour essayer de lever des ambiguïtés comme celles-ci, il est utile de bien distinguer le *langage de la théorie* du *langage des observables*. Le premier, dans lequel sont formulées les théories, lorsqu'elles existent, ou les présupposés qui en tiennent lieu, et aussi les hypothèses, comprend des *concepts* tels que « attitudes », « motivations », « besoin », « représentation », « satisfaction », « aspiration », « idéologie », « rôle », pour citer simplement quelques notions courantes en psychologie sociale et qui interviennent dans l'élaboration ou l'interprétation de nombreuses enquêtes. Interviennent aussi dans le langage de la théorie des *relations* comme « est la cause de... », « entraîne... », ou « influence... », etc.

Ni ces concepts, ni ces relations ne peuvent être exprimés dans le langage des observables, qui comprend des *réponses* à des questions bien déterminées, et des relations telles que des *covariations* ou des *différences* entre paramètres. Des *règles de correspondance* sont nécessaires pour passer d'un langage à l'autre.

Nous verrons plus loin les problèmes que pose la distinction entre les deux langages en ce qui concerne les relations, pour nous en tenir maintenant à ceux qui ont trait aux concepts et aux variables.

On présente souvent le recours aux *définitions opérationnelles* comme un moyen à la fois de résoudre (en le supprimant) le problème de la correspondance entre les deux langages, et d'éviter de travailler avec des concepts purement théoriques ou spéculatifs, sans aucun fondement empirique.

L'expression « définition opérationnelle » a été introduite par le physicien et épistémologue P.W. Bridgeman, à la fin des années vingt, à propos de la construction des concepts en physique. D'inspiration positiviste, il a attiré l'attention sur le fait qu'une notion physique n'est connaissable que par les opérations qui permettent de la mesurer. La température ne devient un concept physique que lorsque, sans se contenter des impressions sensorielles de chaleur, on construit et on standardise des thermomètres. Définir opérationnellement la température, c'est ne pas se poser le problème de son existence en « soi », c'est la *définir* comme ce que mesure le thermomètre. Ainsi, on ne dira pas que c'est la température qui cause l'allongement de la colonne de mercure du thermomètre, mais que par définition, il y a variation de température si la longueur de cette colonne se modifie. Non seulement la notion de température est ainsi définie, en principe, sans ambiguïté, mais on évite également le risque d'explications circulaires comme celles dont se moquait déjà Molière, lorsqu'il ridiculisait les médecins qui prétendaient expliquer par la « vertu dormitive » de l'opium, l'influence qu'a effectivement cette substance sur le sommeil.

Dans cette perspective, le langage de la théorie apparaît simplement comme un moyen, éventuellement le plus commode, de parler des observables, les définitions opérationnelles constituant les règles de passage d'un langage à l'autre. En s'imposant

la discipline d'expliciter les définitions opérationnelles de chaque concept utilisé, on éviterait à la fois des discussions oiseuses sur leur essence, leur « vraie » nature, et on réduirait le risque de parler de notions sans correspondance empirique.

Cette position était déjà celle de Binet lorsque, bien avant Bridgeman, il aurait affirmé son célèbre, et peut-être apocryphe, « l'intelligence, c'est ce que mesure mon test ». Mais il y a entre les deux types d'exemples, psychologique ou physique, des différences essentielles que les psychologues, qui ont accepté avec enthousiasme, surtout aux Etats-Unis, l'idée de définition opérationnelle des concepts, croyant faire comme les physiciens, n'ont pas perçues ou ont négligées.

Cette différence, à notre sens, réside dans l'existence d'une véritable théorie physique, unifiée et de portée générale, dont les quelques théories qu'on rencontre dans les sciences humaines, théories locales qui sont loin de faire l'accord de tous, ne sont que de faibles imitations. L'existence d'une telle théorie en physique a comme conséquence d'empêcher tout arbitraire dans la construction des définitions opérationnelles. En effet, le concept de température, dans le langage de la théorie, présente des propriétés très précises, exprimées par ses relations avec les autres concepts de la théorie, et par les règles de manipulation des énoncés ou des formules où ils apparaissent [1]. On ne peut donc pas décider à volonté, ou en fonction d'intuitions, que telle opération définira adéquatement un certain concept ; il faut encore s'assurer qu'il y a bien correspondance entre les modifications des observables et les manipulations des concepts dans la théorie. D'autre part, c'est encore l'existence d'une théorie qui assure que les différentes définitions opérationnelles possibles d'un même concept sont bien équivalentes. En effet, on peut aussi définir la température ou la mesurer (ce qui est la même chose dans cette perspective) par la pression exercée par un certain volume gazeux ou par la mesure de la différence de potentiel entre les deux extrémités d'un couple thermoélectrique ; d'autres phénomènes, dans lesquels la théorie affirme que la température intervient, pourraient également, bien que peut-être de façon moins commode, servir de définitions opérationnelles, et on pourrait être sûr que toutes seraient équivalentes.

La situation en psychologie [2] est radicalement différente, et par conséquent également le rôle qu'y peuvent jouer les définitions opérationnelles. Sauf dans quelques domaines, comme l'étude de l'apprentissage, où on a cru pouvoir construire de véritables théories [3] avec axiomes et théorèmes, les concepts avec lesquels le psychologue réfléchit, en fonction desquels il formule ses hypothèses, ne sont pas véritablement inclus dans une théorie, au sens où nous avons vu que les concepts physiques le sont, les propriétés ne sont presque jamais définies avec précision, leurs relations avec les autres concepts sont floues. Ceci est d'autant plus vrai des notions intervenant dans les problèmes traités par enquête ; du fait de leur caractère souvent appliqué, ces problèmes ne proviennent pas toujours du développement d'une théorie et de la recherche fondamentale correspondante, mais de questions pratiques dont l'origine est toute différente. Le chercheur est donc fréquemment amené soit à construire des concepts ad

1. Poincaré disait que les axiomes sont « définitions déguisées ».

2. Dans les autres sciences sociales également. Mais, c'est surtout en psychologie que la physique, ou plutôt la représentation idéalisée que s'en font les psychologues, a servi de modèle.

3. Encore que là aussi l'unanimité soit loin d'être faite.

hoc, soit à utiliser des concepts élaborés ailleurs, mais en les isolant de leur contexte, avec tous les riques de glissement de signification que cela peut entraîner. Même lorsqu'on essaie de travailler dans un cadre théorique précis, celui-ci est le plus souvent incomplet, et n'assure donc pas la convergence des différentes définitions opérationnelles qu'on peut concevoir pour un même concept. Il résulte de toutes ces remarques que définir opérationnellement un concept psychologique ou sociologique prend un sens tout différent de celui qu'il a en physique.

Le point essentiel, ici, c'est qu'il y a rarement autre chose que le bon sens pour nous guider dans le choix des opérations à prendre comme définitions de nos concepts (dans le cas des enquêtes, rappelons-le, les opérations sont constituées par les questions que l'on pose). Supposons que nous voulions mesurer l'attitude à l'égard de l'intégration européenne. Tout d'abord, il faut remarquer que c'est notre expérience quotidienne, et non une analyse théorique, qui nous amène à nous poser des problèmes sur cette attitude, à avancer des hypothèses où elle intervient et donc à admettre qu'elle « existe » (nous verrons plus loin quels sont les sens qu'on peut attribuer à ce verbe dans ce contexte). Ce que nous avons pu observer, mais de façon tout à fait informelle, c'est que certaines personnes nous semblent favorables à l'intégration européenne, d'autres défavorables, et ceci à des degrés divers. A partir de cette « constatation » (qui elle-même comprend déjà de nombreux présupposés), c'est encore notre bon sens, ou notre « flair de sociologue », qui nous guideront dans le choix des questions à poser et de leur formulation. D'autres chercheurs pourront peut-être faire d'autres choix, et rien ne nous permet, si ce n'est certaines observations empiriques, d'affirmer que les « mesures » ainsi obtenues seront équivalentes, ni ne nous permettra de prévoir ou d'expliquer les différences entre elles. La définition opérationnelle de l'attitude par les réponses à un ensemble de questions devient une définition arbitraire, fondée uniquement sur le sens commun, sur une expérience en général non systématisée, ou sur une théorie implicite.

Néanmoins, s'imposer de recourir systématiquement à des définitions opérationnelles des concepts avec lesquels on travaille présente le très grand avantage de soumettre ces concepts à un certain type de critique, et d'éliminer ceux à qui, parce qu'ils sont trop vagues ou contradictoires, il est impossible de trouver un correspondant empirique. Un concept qu'il ne s'avère pas possible d'opérationnaliser, même si l'on admet que sa relation avec les observables puisse être complexe ou indirecte, est certainement un concept douteux, qui pose problème. La possibilité d'en donner une définition opérationnelle apparaît donc comme une condition nécessaire, mais non suffisante, de la prise en considération d'une notion [1].

En effet, se contenter d'une telle définition, et affirmer que lorsqu'on parle d'une certaine attitude, on ne parle de rien d'autre que des réponses à un certain nombre de

1. Sur le plan épistémologique, on peut même douter que ce soit une condition nécessaire, c'est-à-dire qu'il soit effectivement possible de donner une définition opérationnelle de tout concept jugé scientifique. Mais discuter ce point ici nous entraînerait trop loin. Il faut toutefois remarquer que, dans les sciences sociales tout au moins, certains concepts très généraux ne peuvent pas être directement définis de façon opérationnelle. On peut opérationnaliser l'attitude à l'égard d'un certain objet, on ne peut pas donner de définition opérationnelle de la notion d'attitude en général. Il en va de même pour des concepts tels que aptitude, rôle, représentation, qui apparaissent plutôt comme des classes de notions définies empiriquement. Il ne semble pas que ce problème ait son équivalent en physique.

questions bien déterminées, c'est se prémunir à l'avance, sous un prétexte de pureté méthodologique, contre toute critique portant sur le contenu de l'instrument de mesure utilisé, autrement dit sur le choix des questions. S'en tenir rigoureusement, et jusqu'au bout, à cette correspondance est, en pratique, impossible : ce ne sont pas, en effet, ces questions particulières qui nous intéressent, mais ce qu'elles sont censées représenter ; si ce n'était pas le cas, on parlerait directement de ces questions, pas d'une attitude. On a donc besoin, ne serait-ce qu'au moment de tirer les conclusions de l'enquête, d'énoncer des propositions qui porteront sur l'attitude elle-même, mais en revenant à son sens intuitif initial, avec toutes ses connotations, sens qui ne se limite pas à un ensemble particulier de questions. La prudence méthodologique ne peut pas être maintenue jusqu'au bout.

Le problème est un peu différent si, au lieu d'étudier une attitude particulière, on s'intéresse à des mécanismes généraux, par exemple aux processus d'influence et aux changements d'attitude. L'ensemble particulier des questions que choisira le chercheur, puisqu'il lui faut bien concrétiser la notion d'attitude, ne l'intéressera pas par lui-même. Il n'aura donc pas à se demander si son opérationnalisation est adéquate. Mais dans la plupart des enquêtes, c'est au contenu lui-même qu'on s'intéresse. Que les notions utilisées proviennent d'une théorie, ou, ce qui est plus fréquent, de l'expérience quotidienne du chercheur, ce n'est que par rapport à cet arrière-plan que la définition opérationnelle prend son sens.

Prendre trop au sérieux la notion de définition opérationnelle, et donc se refuser à parler d'une attitude distincte des questions posées, peut avoir d'autres conséquences négatives. Par exemple, devant un résultat empirique obtenu, ce point de vue ne conduira pas à se demander s'il est propre à l'opérationnalisation particulière avec laquelle on a travaillé, ou si c'est bien une propriété de l'attitude en question, donc généralisable aux autres opérationnalisations possibles. Si avec une autre, on obtient un résultat différent, on devra admettre qu'il s'agissait de deux attitudes différentes. En revanche, c'est une question qu'on sera tout naturellement amené à se poser si on admet l'existence de l'attitude préalablement à toute mesure ; à ce moment chaque ensemble de questions ne constituera que *l'un des indicateurs possibles* de cette attitude, indicateur dont on pourra se demander s'il est le meilleur et comment il est biaisé par rapport à d'autres indicateurs possibles.

Cela débouche sur une méthodologie plus exigeante, peut-être difficile à mettre en œuvre. Elle impose, en premier lieu, qu'on explicite le fondement de notre croyance dans l'existence du concept en question. Cela est, en principe, aisé lorsqu'il intervient dans une théorie, mais beaucoup plus difficile lorsque ce fondement est enraciné dans une expérience quotidienne, plus ou moins éclairée par des connaissances psychologiques ou sociologiques, et donc dans une sociologie ou une psychologie « spontanées ». Il faudrait pour cela commencer par admettre l'existence de ce lien entre les pratiques des chercheurs en sciences sociales et leur vie quotidienne, ce qui non seulement est difficile lorsqu'on cherche à dépasser les affirmations de principe, mais remet aussi en cause une certaine image « scientifique » de ces disciplines [1].

1. A l'intérieur de la sociologie anglaise et américaine, le courant « ethnométhodologique » veut précisément prendre comme objet ces relations entre la vie quotidienne et les pratiques des sociologues. Mais, jusqu'à présent, ils n'ont guère obtenu de

En second lieu, raisonner et parler en termes d'indicateurs amène à ne pas, en principe, se contenter d'une seule mesure d'un concept, mais à en proposer plusieurs, autant que possible faisant appel à des techniques très différentes. Pour mesurer une attitude, par exemple, on introduira donc, dans un même questionnaire, des questions d'opinion de formes très différentes, des questions sur des comportements personnels, des jugements sur des comportements d'autres personnes, des questions sur des comportements probables dans des situations hypothétiques, etc. On cherchera aussi à recouper les résultats de l'enquête par ceux obtenus par d'autres méthodes. En toute rigueur on ne pourra affirmer qu'une attitude « existe » que lorsqu'on aura montré empiriquement que toutes ces « opérations », tous ces indicateurs, convergent, sont bien cohérents entre eux. L'attitude, ou plus généralement toute variable sous-jacente, ce sera ce qu'il y a de commun à ces différentes opérationnalisations. Certes, de cette manière, on n'évite pas que chacun des indicateurs élémentaires n'ait été choisi en fonction des présupposés des chercheurs dont nous parlions, de leur conception intuitive de ce que devrait être cette variable. Mais la vérification de leur cohérence constitue une précaution supplémentaire.

7.2.1.2. La construction des échelles d'attitude

a) Généralités

Lorsqu'une réponse, ou un ensemble de réponses, sont utilisés comme indicateurs d'une variable sous-jacente, il faut s'assurer que les propriétés de l'indicateur (ou du moins les propriétés que l'on retiendra) correspondent bien aux propriétés de cette variable, telles qu'on peut les supposer ou les postuler. Construire un indicateur, c'est donc construire une variable ayant les propriétés qu'il doit représenter.

C'est un problème qui est habituellement traité comme un problème de *mesure*. En psychologie sociale, ce sont les méthodes de mesure des attitudes qui ont surtout été développées, bien que de nombreuses autres variables posent les mêmes problèmes de principe, même si elles exigent le recours à des techniques différentes. Nous verrons plus loin quelques exemples de tels indicateurs. Pour le moment, et tant que nous restons dans les généralités, ce sont les attitudes qui nous serviront d'exemple. L'imprécision habituelle dans la définition de cette notion permet d'ailleurs de recouvrir un champ suffisamment large.

Mesurer, c'est à la fois confondre et distinguer, c'est négliger certaines différences et organiser les autres [1]. C'est accepter que, malgré tout ce qui les distingue, des individus puissent être placés dans la même catégorie ou comparés entre eux ; c'est aussi admettre que la relation d'ordre, et éventuellement d'autres relations qui viendraient s'y ajouter, entre les catégories, décrivent adéquatement le phénomène étudié.

---

résultats utilisables à des fins méthodologiques, malgré l'intérêt de leur approche critique et la valeur de leurs travaux dans d'autres domaines.

1. Contrairement à la définition mathématique, mais en accord avec un usage courant dans les sciences sociales, nous appellerons ici *mesure* toute structure comportant au moins une relation de *pré-ordre total* . Dans le langage des niveaux de mesure, familier aux psychologues, il s'agit des échelles ordinales, d'intervalle et de rapport.

Par exemple, on pourra décider de « mesurer » une attitude par le nombre de questions auxquelles chaque individu interrogé aura répondu « oui ». Cela signifie qu'on acceptera de confondre, de placer dans la même catégorie, des personnes qui auront donné des réponses très différentes les unes des autres, peut-être même qui n'auront donné aucune réponse commune. D'autre part, cela signifie également que l'ordre ainsi obtenu, celui des nombres de réponses, représente bien l'attitude en question, c'est-à-dire que ceux qui ont donné plusieurs réponses « oui » peuvent être considérés comme « plus », du point de vue de l'attitude, que ceux qui en ont donné un plus petit nombre.

Bien que cette manière de mesurer soit simple et apparemment raisonnable, ce n'est pas la seule possible. On pourra par exemple attribuer à chaque réponse « oui » un poids variant d'une question à l'autre. Quelle que soit la manière dont on aura décidé de ces poids, cela aura comme conséquence qu'on devra classer dans la même catégorie, si cela se produit, une personne ayant répondu « oui » à une seule question, mais dont le poids est très élevé, et une autre qui aura répondu « oui » à plusieurs questions de poids faible. On pourra aussi ne considérer comme significative que la réponse la plus extrême donnée par une personne, indépendamment de ses autres réponses. Toutes les personnes placées dans la même catégorie auront alors en commun d'avoir répondu « oui » à une même question et « non » aux questions plus extrêmes, en négligeant leurs réponses aux autres questions.

Ces exemples (où nous n'avons pas abordé le problème de savoir comment on attribuerait leurs poids aux questions, ni comment on jugerait leur degré d'extrémisme) montrent que la constitution d'un instrument de mesure, dans le cas particulier une échelle d'attitude, n'est pas évident, même si l'on estime que les questions posées ont été bien choisies.

Lorsque la mesure est effectuée au moyen d'une seule question fermée, le problème est formellement très simple ; si les réponses qu'on propose sont peu nombreuses, et ordonnées sans ambiguïté (exprimant par exemple différents degrés d'accord avec l'opinion énoncée dans la question), la mesure est immédiatement fournie par la réponse choisie elle-même : ceux qui ont donné la même réponse sont ipso facto confondus, et l'ordre des catégories est donné par celui des réponses. Certes, cela suppose que chacun ait compris la question et les réponses de la même manière, et qu'on admette qu'il n'est pas possible de choisir la même réponse pour des raisons trop différentes. Mais cela renvoie à une critique des formulations elles-mêmes, et non à une procédure formelle de mesure.

Toutefois, lorsqu'on veut saisir une variable de portée générale, il est rare qu'on se contente d'une seule question. Les raisons en sont multiples.

En premier lieu, mais c'est loin d'être la raison la plus importante, une seule question ne donne qu'un nombre restreint de catégories, autant que de réponses, et on souhaite parfois faire des distinctions plus fines. On pourrait, dans ce cas, augmenter le nombre des réponses proposées, par exemple en choisissant une échelle numérique au lieu d'une échelle verbale. Mais on sait que ces précisions deviennent très vite illusoires, et qu'il est inutile de demander aux personnes interrogées des distinctions qui ne correspondent pas à grand-chose pour elles.

En second lieu, on peut craindre que la réponse à une question unique ne soit trop in-

fluencée par les caractéristiques propres de cette question (thème particulier abordé, vocabulaire, etc.). En posant plusieurs questions, on peut penser que de tels biais liés à la spécificité de chaque question se compenseront.

Enfin, et c'est là une raison majeure, lorsqu'on pose plusieurs questions, on peut *vérifier si elles présentent une cohérence suffisante pour pouvoir être acceptées comme des indicateurs de la même variable sous-jacente,* ce qui, seul, peut justifier de combiner les différentes réponses en un indicateur unique, et donner à la validité de celui-ci une vraisemblance suffisante.

Les techniques de construction d'échelles, quels que soient leurs fondements, ne se limitent donc pas à des procédures de combinaisons de réponses pour constituer un indicateur unique, mais comprennent toutes une phase de vérification de cohérence, qui justifiera cette procédure.

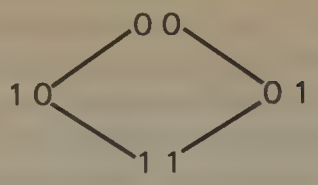
Mesurer une attitude, c'est se donner la possibilité d'ordonner tous les individus étudiés selon qu'ils sont plus ou moins favorables à un certain objet, ou, plus généralement, qu'ils présentent plus ou moins une certaine tendance. Construire une échelle d'attitude à l'égard de l'intégration européenne, par exemple, c'est poser des questions dont les réponses permettent d'ordonner les sujets selon qu'ils lui sont plus ou moins favorables. Mais on peut aussi considérer des attitudes de portée plus générale comme l'attitude autoritaire, par exemple, ou la croyance religieuse, etc. Formellement, la démarche reste la même.

Pour comprendre le principe de cette démarche, plaçons-nous dans le cas où nous avons posé une série de questions que nous supposons se rapporter à la même attitude et dont nous attendons qu'elles constituent l'échelle voulue. Pour simplifier, nous admettons pour le moment qu'à chaque question ne correspondent que deux réponses, l'une « favorable », que nous représentons par 1, l'autre « défavorable », notée 0.

Si, pour commencer, nous ne considérons qu'une seule question, ces deux réponses définiront deux groupes de sujets. Nous avons vu les insuffisances d'une classification aussi sommaire. Passons à deux questions ; elles définissent à priori quatre combinaisons de réponses, donc quatre groupes de sujets :

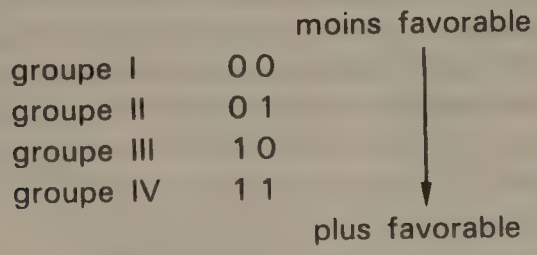
	Questions	
	A	B
groupe I	0	0
groupe II	1	0
groupe III	0	1
groupe IV	1	1

Le groupe I peut être considéré comme le plus défavorable, le groupe IV comme le plus favorable, les deux autres étant intermédiaires, sans qu'il soit possible, avec la seule hypothèse que nous avons faite (que l'ordre 0-1 a la même signification dans les deux questions), de les ordonner. Tout ce que nous avons, c'est un ordre partiel :



Pour obtenir un ordre total, on peut envisager plusieurs démarches qui toutes font appel à des postulats supplémentaires. La technique la plus simple consiste à regrouper les sujets ayant donné les combinaisons de réponses 0 1 et 1 0 en une seule catégorie intermédiaire. En mettant provisoirement entre parenthèses le problème de la légitimité de la composition des réponses à deux questions différentes, remarquons que la constitution de ce groupe intermédiaire unique postule que les deux questions sont équivalentes, et que le fait de répondre favorablement à l'une seulement, et non à l'autre, est purement contingent, et sans signification en ce qui concerne l'attitude sous-jacente que l'on cherche à mesurer. Cette manière de faire a le mérite de la simplicité, mais on voit qu'elle s'appuie sur un postulat très fort.

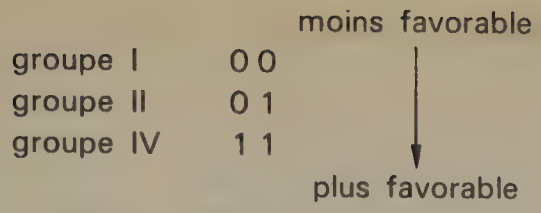
On peut aussi, pour éviter de postuler l'équivalence entre les questions, attribuer à chaque réponse un « poids » différent, et caractériser chaque sujet par la somme des poids des réponses qu'il a données. (Dans l'exemple précédent, chaque réponse favorable avait le poids 1, chaque réponse défavorable le poids 0.) Par exemple, on pourra estimer que A exprime une position très favorable, et B une position moins favorable. On attribuera alors à la réponse 1 à A un poids supérieur à la réponse 1 à B. On obtiendra alors l'ordre suivant :



Remarquons que, quels que soient les poids choisis, l'ordre des groupes restera toujours compatible avec l'ordre partiel : il est impossible de trouver un système de poids additifs (linéaire) qui ne désigne pas le groupe I et le groupe IV respectivement comme le moins et le plus favorable, les deux autres étant intermédiaires.

L'attribution de poids aux réponses et la composition additive de ces poids, caractérisent deux des méthodes que nous allons exposer, celle de Likert et celle de Thurstone, qui diffèrent par le mode de détermination des poids et par le type de justification donné de la composition de ceux-ci.

Supposons maintenant qu'un des deux groupes intermédiaires, par exemple le groupe III, ne soit pas représenté ; les trois groupes qui restent sont alors ordonnés sans ambiguïté :



Cette possibilité, qui peut être généralisée au cas d'un nombre quelconque de questions, est à la base des techniques dites d'analyse hiérarchique, la troisième méthode que nous exposerons.

b) La méthode de Likert

On suppose que l'attitude existe, c'est-à-dire que les sujets sont ordonnés selon qu'ils sont plus ou moins favorables à l'objet étudié. Si les questions se rapportent bien à l'attitude en question, les probabilités pour qu'un sujet donne la réponse favorable à chaque question particulière seront d'autant plus fortes qu'il sera lui-même plus favorable (cf. fig. 1). La courbe correspondant à la question A donne les probabilités de réponse favorable à une question « modérée », réponse que même des sujets peu favorables peuvent donner. La question D, en revanche, ne suscite de réponse favorable que de la part de sujets globalement très favorables. Les questions B et C sont intermédiaires.

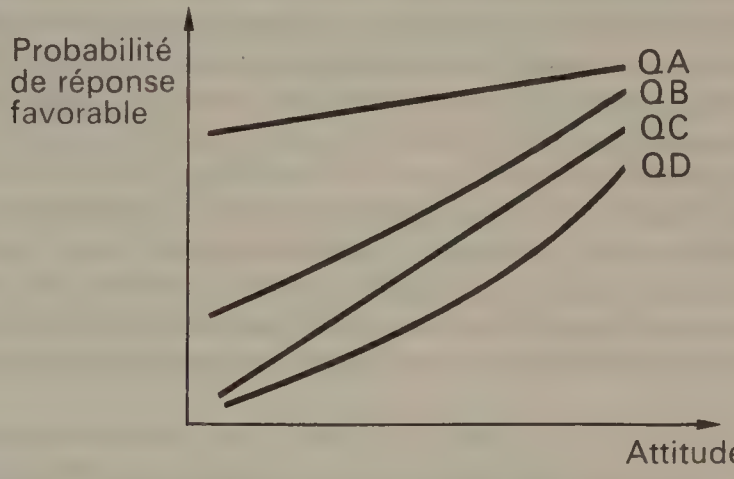


Fig. 1

Ce modèle exclut la possibilité d'autres courbes comme celles de la fig. 2. Par exemple, la courbe E, qui traduit une proportion constante de réponses favorables, indique que ces réponses sont indépendantes de l'attitude. La question portait donc sur autre chose, et c'est par erreur qu'elle a été introduite dans l'échelle. La courbe F correspond simplement à une erreur de codage. C'est l'autre réponse qui exprime une position favorable, cela peut se corriger facilement. Quant à la courbe G, c'est celle qu'on obtiendrait avec une réponse exprimant une position fondamentalement « moyenne », refusée aussi bien par les personnes nettement favorables que par celles qui sont nettement défavorables. Nous avons vu des exemples de telles questions en parlant de la formulation des questions d'opinion (cf. 4.5.3.3.). Elle n'est pas nécessairement « mauvaise », mais elle ne peut pas être combinée avec les autres.

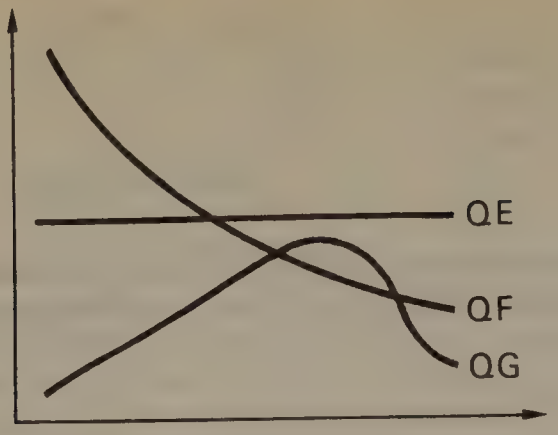


Fig. 2

Le modèle suppose donc que *la probabilité de la réponse favorable à une question croît de façon monotone en fonction de l'attitude sous-jacente des sujets.* Devant un ensemble de questions qu'on suppose se rapporter toutes à la même attitude, il faut donc vérifier que les réponses recueillies sont bien conformes au modèle. C'est cela qui constituera le critère de cohérence nécessaire pour admettre que les différentes questions peuvent intervenir dans la même échelle.

Ce modèle justifie également une composition additive des réponses. En effet, on voit qu'une personne très favorable a une forte probabilité de donner un grand nombre de réponses favorables. Inversement, une personne ayant donné plusieurs réponses favorables a une probabilité élevée d'avoir une attitude favorable. Le nombre de réponses favorables, ou éventuellement une combinaison plus complexe, constituera donc un bon indicateur de l'attitude.

Si nous disposions déjà d'une mesure indépendante de la variable sous-jacente, la vérification de l'adéquation des données recueillies au modèle serait immédiate. On n'aurait qu'à tracer les courbes en question, et à vérifier qu'elles sont effectivement monotones, en rejetant les questions donnant des courbes d'une autre forme. Mais c'est supposer le problème résolu, puisque c'est précisément cette mesure que nous cherchons.

La méthode proposée par Likert consiste à utiliser une mesure provisoire de l'attitude, en prenant comme estimation le nombre de réponses favorables. On constitue ainsi des groupes de sujets ayant donné le même nombre de réponses favorables. Dans chaque groupe, on peut calculer la proportion de sujets ayant donné une réponse favorable à une question particulière. Si cette proportion croît régulièrement en fonction de cette estimation de l'attitude, on aura vérifié, ne serait-ce que de façon grossière, que la probabilité de réponse favorable à cette question varie bien de façon monotone avec l'attitude supposée (fig. 3).

Cette manière de procéder semble circulaire, puisque la réponse à la question étudiée intervient à la fois en abscisse (elle a été comptée dans le nombre de réponses favorables) et en ordonnée. En fait, l'erreur qu'on commet ainsi est faible, surtout si le nombre de questions est élevé. S'il est trop faible, inférieur à 5 ou 6, on pourra prendre comme estimation de l'attitude globale le nombre de réponses favorables à toutes les

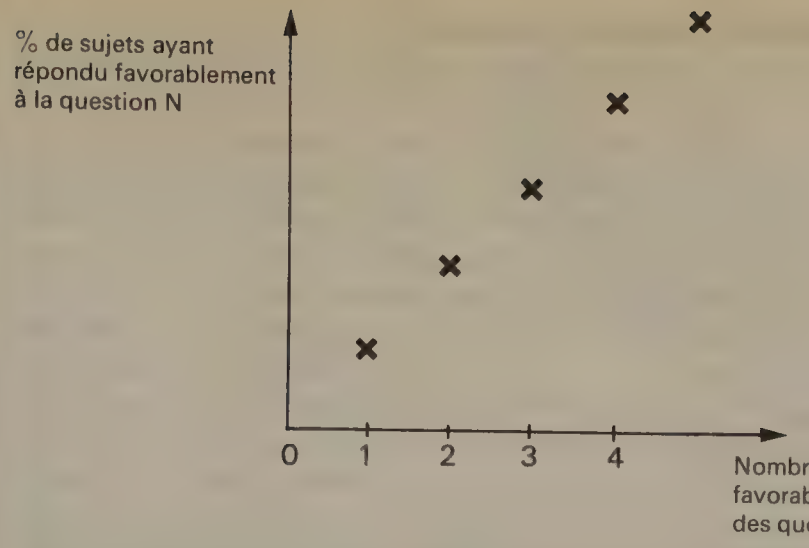


Fig. 3

questions, sauf celle dont on vérifie la cohérence avec les autres. Mais cela complique beaucoup la procédure, puisqu'il faut calculer pour chaque question une nouvelle estimation de l'attitude.

La monotonie des courbes ainsi obtenues n'est jamais parfaite. Pour en avoir un critère, on pourra calculer un coefficient de corrélation bisérial, entre une variable métrique (le nombre de réponses favorables à l'ensemble des questions) et une variable dichotomique (la réponse à une question particulière). Un coefficient significatif permettra d'accepter l'hypothèse de monotonie, donc d'adéquation de la question au modèle.

Cela oblige à éliminer certaines questions. Il faudrait alors répéter l'opération avec celles qu'on aura retenues, jusqu'à ce que toutes les questions donnent une courbe monotone.

Lorsqu'on en a la possibilité, on peut être plus exigeant, et éliminer des questions donnant des courbes comme celles de la fig. 4, bien qu'elles soient monotones. Cet enchevêtrement indique que la relation de chaque question avec les autres n'est pas de

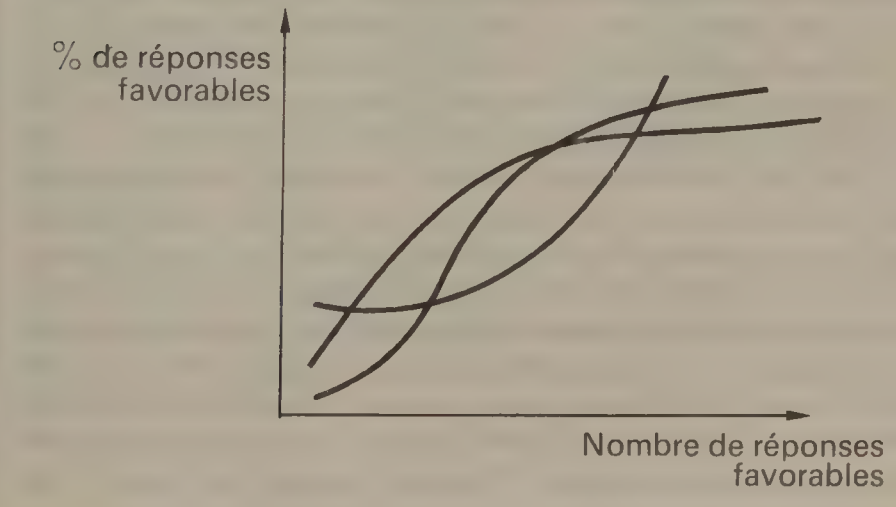


Fig. 4

Fig. 5

même nature. On ne sait plus très bien ce que signifient les réponses à de telles questions. C'est pourquoi on préfère les éliminer, et ne conserver que des questions dont les courbes sont à peu près de même forme et ne se coupent pas (fig. 5).

Par ce premier contrôle, nous avons vérifié que les différentes questions retenues sont bien cohérentes entre elles, donc qu'il est légitime de les considérer comme se rapportant à une même variable sous-jacente. Lorsque le nombre de questions rejetées est faible, on peut penser que leur élimination n'a guère de conséquences. S'il est plus important, il faut se demander si les questions retenues couvrent bien le champ prévu au départ. Pour cela, il faut essayer d'expliquer pourquoi certaines questions ont satisfait le critère et pas les autres. Si on ne le fait pas, on risque de commettre des erreurs graves concernant la signification de l'attitude mesurée par les questions restantes. Supposons, pour prendre un exemple simpliste, que nous voulions construire une échelle d'attitude à l'égard de l'intégration européenne. Nous poserons pour cela un grand nombre de questions portant sur différents aspects de l'Europe : politique, économique, monétaire, culturel, etc. On parlera aussi bien de fixation des prix agricoles, que de la libre circulation des travailleurs ou d'élections directes d'un parlement européen. Nous posons toutes ces questions à un échantillon de personnes, et vérifions qu'elles présentent bien la cohérence définie par le modèle ; cela nous amènera peut-être à éliminer certaines questions. C'est alors qu'il faudra revenir à leur contenu ; si l'on constate que toutes les questions retenues se rapportent, disons, aux aspects économiques de l'Europe, il faudra prendre garde au fait que l'échelle ainsi construite ne portera pas, comme on le voulait au départ, sur l'intégration européenne en général, mais uniquement sur l'un de ses aspects : sa signification a changé, et il faudra en tenir compte chaque fois que l'on cherchera à interpréter des résultats où cette échelle intervient.

Cette nécessité d'éliminer certaines questions peut aussi constituer une information positive. Dans certains cas, elle indiquera que le champ posé au départ n'était en fait pas homogène ; dans le cas particulier, cela voudra dire qu'il n'y a pas, dans la population, d'attitude globale à l'égard de l'Europe, mais seulement des attitudes à l'égard de certains aspects particuliers, ces attitudes partielles n'étant pas coordonnées entre elles. Ou alors, ce même résultat peut être attribué à l'existence, dans la population étudiée,

de sous-groupes distincts présentant des cohérences différentes. Il faudrait alors faire des hypothèses sur les caractéristiques de ces sous-groupes et étudier l'attitude en question séparément dans chacun d'eux.

Enfin, il ne faut pas exclure que l'absence de cohérence entre certaines questions ne soit due qu'à des raisons tenant à la forme des questions, et non à leur contenu. Nous avons insisté, ci-dessus, sur l'intérêt de travailler avec des indicateurs élémentaires de nature très différente. Les raisons de principe que nous avons avancées pour cela sont sérieuses, mais il faut prendre garde qu'on augmente ainsi les risques de non-cohérence apparente, non-cohérence qui ne serait pas due au contenu des questions lui-même, mais seulement à leur forme.

Lorsque ces vérifications ont été effectuées, il reste un ensemble de questions cohérentes. Nous avons vu qu'il était ainsi justifié de compter le nombre de réponses favorables, et de prendre ce nombre comme indicateur de l'attitude. Mais on peut aussi raffiner cette mesure, en ne se contentant pas d'attribuer des poids égaux à chaque réponse, mais en tenant compte du fait que certaines réponses expriment une position très favorable, et d'autres une position qui l'est beaucoup moins.

Pour choisir les poids, on partira du principe qu'*une réponse est d'autant plus significative, donc doit avoir un poids d'autant plus grand, qu'elle est donnée par peu de personnes*. Si la réponse favorable à une certaine question est donnée par une forte majorité, on sait qu'elle aura été donnée aussi bien par des personnes très favorables que par des personnes modérées, voire défavorables. A elle seule elle n'apportera donc pas beaucoup d'information, et on lui attribuera un poids faible. En revanche, l'autre réponse à cette même question n'aura été donnée que par peu de personnes, nécessaire-

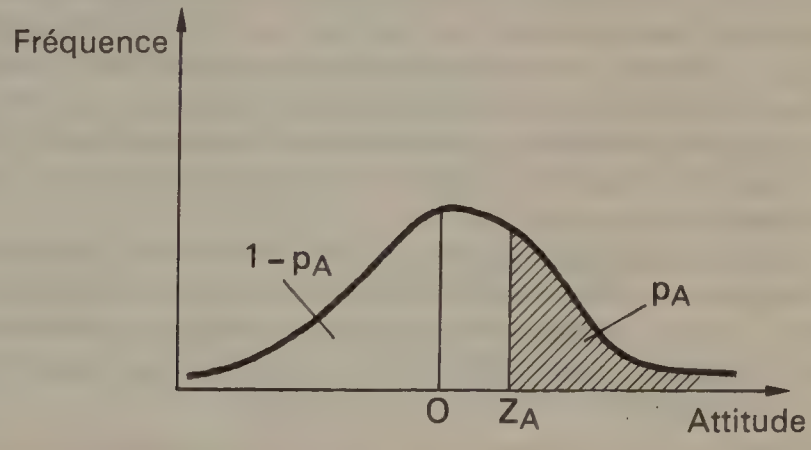


Fig. 6

ment des personnes défavorables [1]. On lui donnera alors un poids élevé en valeur absolue, mais de signe contraire du poids de l'autre réponse.

Pratiquement, on procède de la manière suivante. On postule que la distribution des sujets sur la variable sous-jacente est gaussienne, de moyenne 0 et d'écart-type 1 (fig.

1. Si ce n'était pas le cas, la courbe correspondant à cette question n'aurait pas été monotone croissante.

6). Cela revient à fixer la métrique de l'attitude. En effet, celle-ci n'est à priori pas déterminée. Fixer la forme de la distribution, c'est adapter tous les intervalles sur l'échelle, donc choisir une métrique particulière.

Chaque question divise cette distribution en deux parties, selon la réponse, favorable ou défavorable, qu'ont donnée les sujets. Si une question A suscite une proportion p_A de réponses favorables, et donc 1-p_A de réponses défavorables, il est possible de trouver (en consultant simplement la table de la loi normale réduite) une valeur Z_A qui divise la distribution dans les mêmes proportions que la question. Certes, nous savons bien que tous les sujets ayant répondu favorablement à A ne se situent pas nécessairement tous du même côté de Z_A, mais, en l'absence d'autres informations, on l'admettra, comme approximation. On leur attribuera une note égale à la médiane de l'intervalle $Z_{A1} + \infty$. Cette valeur, Z_{A1}, se trouve également dans la table de la loi normale réduite, c'est celle qui correspond à $p_A/2$. Nous pouvons faire de même pour la réponse négative, et attribuer à tous ceux qui l'ont donnée le poids Z_{A0}, correspondant à la proportion 1 - $P_A/2$.

Cette opération peut être répétée pour chaque question : à chacune correspondront deux valeurs, une pour chaque réponse. Chaque valeur peut être considérée comme une évaluation de la position du sujet sur l'échelle, estimée à partir d'une seule de ses réponses. Cela amène à prendre comme estimation finale la *moyenne* des estimations faites à partir de chaque question. Comme tous les sujets ont répondu au même nombre de questions, faire simplement la somme des poids, ou leur moyenne, est équivalent et revient à un changement d'unité.

La composition additive des poids de chaque réponse peut ainsi se justifier de deux manières. Premièrement, en admettant qu'un individu a d'autant plus de chances d'être « réellement » favorable qu'il a donné plus de réponses favorables, et des réponses favorables plus rares. Deuxièmement, en considérant que chaque réponse nous donne une estimation de la position du sujet, une meilleure estimation étant donnée par la moyenne de celles-ci. En fait, ces deux justifications sont équivalentes, et non seulement par le fait qu'elles aboutissent à la même technique. Dans chaque cas, on infère la position du sujet à partir de ses réponses, en s'appuyant sur le modèle de croissance monotone des probabilités de réponse.

Notons que si le modèle est nécessaire pour justifier la procédure, celle-ci peut s'appliquer à n'importe quelles données, ce qui est dangereux. En effet, quelles que soient les questions posées, et leurs relations, il est toujours possible d'attribuer un poids à chaque réponse et d'en faire la somme, même si cette opération est dénuée de toute signification, si ce n'est d'exprimer les a priori du chercheur.

En principe, le modèle sous-jacent que nous avons exposé est probabiliste. En fait, le traitement lui-même ne l'est pas [1], le lien avec le modèle restant, comme nous l'avons vu, très intuitif. La composition additive des poids ne tient pas compte explicitement de la possibilité de fluctuation des réponses des sujets, donc d'une incertitude sur ce point. Mais par l'idée de compensation que la méthode implique, on laisse en fait la possibilité de tenir compte de ces variations.

---

1. Si on tire les conséquences du modèle, on retrouve l'analyse de structure latente.

L'attribution à chaque individu d'un score par addition des poids de ses réponses a comme conséquence qu'un même score peut être obtenu de plusieurs manières différentes. Ainsi, il est possible de confondre un sujet ayant donné une seule réponse très favorable, et un autre qui aura donné plusieurs réponses modérées. Il y a donc bien, sous-jacente, l'idée d'une compensation possible entre les réponses. On peut d'ailleurs se demander quelle est la signification, du point de vue de la cohérence des réponses, de l'existence de configurations de réponses aussi différentes. Pourquoi une personne qui a donné une réponse exprimant une position très favorable, donc rare, ne répond-elle pas également de façon favorable à des questions plus modérées ? Cela peut être parce que ayant une possibilité extrême, elle refuse les opinions moyennes. Si c'est le cas, alors la placer dans la même catégorie que ceux qui ont donné uniquement des réponses modérées constitue une erreur. En fait, la forme monotone des courbes indique que de tels cas, sans être impossibles, sont très rares, et l'erreur qu'on commet ainsi n'est pas grave, au niveau de l'ensemble de l'échantillon, même si elle l'est pour un individu particulier. Le fait que les courbes de probabilité soient monotones indique, nous l'avons vu, qu'il n'y a pas de questions qui expriment des positions intrinsèquement modérées. Les sujets extrêmes, dans leur majorité, approuvent également les opinions plus moyennes.

Il arrive souvent que, pour des raisons diverses, qui ont été mentionnées (cf. 4.5.3.2.), on estime que proposer deux réponses n'est pas suffisant. On se trouve donc avoir plusieurs réponses ordonnées de la moins favorable à la plus favorable. La procédure pratique est un peu plus compliquée, mais le raisonnement reste exactement le même.

Pour la première étape, la vérification de la cohérence des questions par la monotonie des courbes de fréquence, le plus simple consiste à regrouper les réponses en deux catégories de manière à se ramener au cas précédent. On perd ainsi un peu d'information, mais le risque d'introduire des erreurs importantes est minime.

Le calcul du poids de chaque réponse se fait selon les mêmes principes que dans le cas dichotomique. Chaque question divise la distribution des individus en autant de groupes qu'il y a de réponses possibles (cf. fig. 7). Connaissant la proportion des sujets dans chaque catégorie (c'est la proportion de ceux qui ont donné la réponse correspondante) on peut, en se référant à la table de la loi normale réduite, en calculer la valeur médiane. Un exemple rendra la procédure plus claire.

A une question où les sujets avaient le choix entre 4 réponses, les proportions de ceux qui ont donné chacune d'elles ont été :

R1	25 %
R2	30 %
R3	35 %
R4	10 %

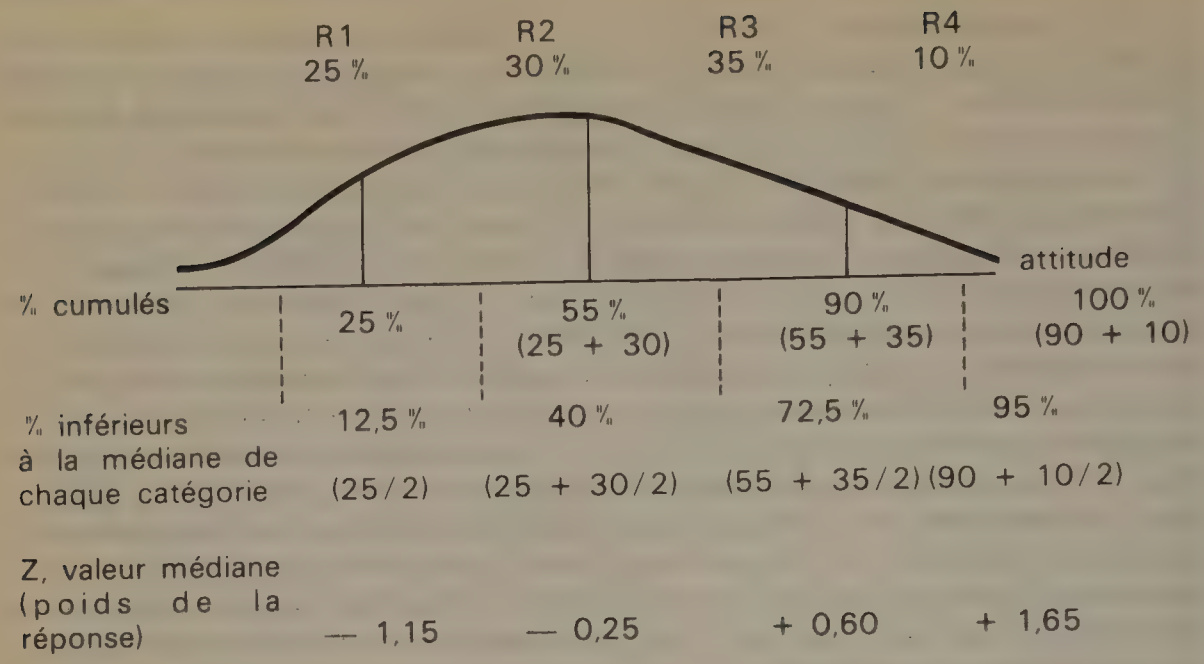


Fig. 7

Le passage de l'avant-dernière ligne (% inférieur à la médiane) à la dernière (7) se fait par lecture de la table de la loi normale réduite, en passant des pourcentages cumulés à la valeur Z correspondante.

c) La méthode de Thurstone

C'est une méthode de pondération, comme celle de Likert. Les poids n'y sont toutefois pas déterminés par la fréquence de chacune des réponses, mais, au cours d'une phase préalable à l'enquête proprement dite, par le recours à des *juges*. Cette méthode s'applique bien dans le cas classique où l'on demande à des sujets de se déclarer d'accord ou non avec des opinions qui leur sont proposées, mais on peut sans difficulté l'étendre à d'autres matériels, par exemple à des comportements considérés comme des indicateurs d'attitude ou de préférences. On utilise également la méthode, limitée à sa première phase, lorsqu'on veut étudier la structure des jugements d'une population considérée comme homogène.

Lors de la première phase, la tâche demandée aux juges consiste à comparer les différents items (par exemple les énoncés d'opinion) du point de vue de l'attitude qu'ils traduisent. On leur présentera souvent ces items par paires (méthode dite du *jugement comparatif*), en demandant lequel des deux exprime l'opinion la plus favorable, indépendamment de l'opinion personnelle de chaque juge. En principe, ceux-ci doivent faire abstraction de leur propre position, et adopter une attitude « objective ». On leur demandera « Laquelle de ces deux opinions est la plus favorable à... ? » et non « Avec laquelle de ces deux opinions êtes-vous le plus d'accord ? ». Sous-jacent à la méthode se trouve le postulat que deux personnes peuvent donner la même réponse à la première question même si elles diffèrent à la deuxième.

Lorsqu'il y a beaucoup d'items, le nombre de paires à présenter devient vite trop grand en pratique. On demande alors plutôt aux juges de classer les items dans un certain nombre (généralement entre 7 et 13) de catégories ordonnées. Certaines méthodes fixent même le nombre d'items à placer dans chaque catégorie, ce qui impose la forme de leur distribution. Cette technique, avec toutes ses variantes, constitue ce qu'on appelle la méthode du *jugement catégoriel.*

Les deux méthodes (jugement comparatif et jugement catégoriel) ont été élaborées par Thurstone à la fin des années vingt. Elles sont fondées exactement sur les mêmes postulats relatifs aux mécanismes du jugement.

A partir des jugements obtenus au cours de la première phase, on vérifie l'homogénéité des items, et on déduit le poids de chacun. Au cours de la seconde phase, ces items seront introduits dans un questionnaire et proposés à un échantillon de la population étudiée, à qui on demande alors d'exprimer son accord ou son désaccord. Chaque sujet interrogé sera ensuite caractérisé par la somme des poids des opinions qu'il aura approuvées.

Le modèle utilisé pour la détermination des poids à partir des jugements est un modèle probabiliste. On pose, comme dans la méthode de Likert, l'existence de l'attitude, variable unidimensionnelle ; mais, dans le cas présent, ce sont les opinions qui sont situées sur cette échelle, et non les individus (les juges), puisque ceux-ci sont censés être « objectifs » et apprécier seulement la position relative des items. Mais, et c'est en cela que le modèle est probabiliste, la valeur d'un item selon cette dimension n'est pas fixe : elle est susceptible de fluctuer de façon aléatoire autour d'une valeur moyenne (celle qu'on cherche à déterminer) selon une distribution que l'on postule gaussienne. Les raisons de cette fluctuation ne sont pas explicitées ; on peut admettre qu'il s'agit de variations dans l'état du sujet qui, d'un moment à l'autre, ne percevra pas une même opinion de la même manière : tantôt elle lui apparaîtra très favorable, à d'autres moments elle lui semblera très modérée, selon son humeur, les influences récentes qu'il aura subies, etc. Une autre opinion sera caractérisée par une autre moyenne, éventuellement une autre variance, la différence de variance exprimant la possibilité que certains items soient plus stables que d'autres (fig. 8).

Lorsqu'on présente simultanément à un juge les items A et B, en lui demandant lequel exprime l'attitude la plus favorable, il les comparera selon les valeurs instan-

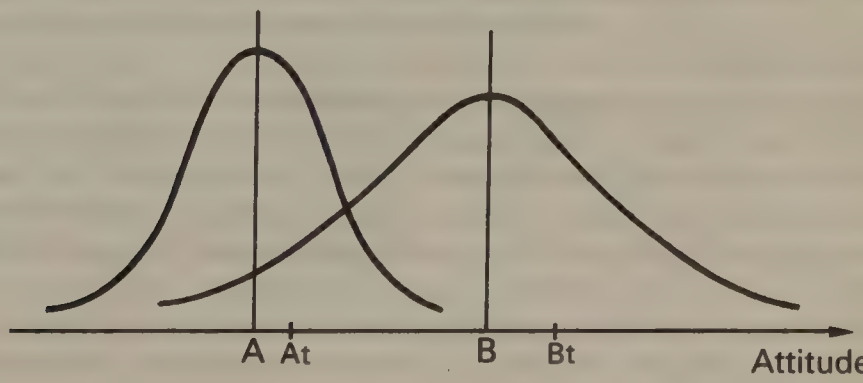


Fig. 8

tanées A_t et B_t qu'ils auront pour lui à ce moment. Dans le cas de la figure 8, B_t sera plus souvent supérieur à A_t que l'inverse, et le juge dira donc plus souvent que B exprime une attitude plus favorable que A. Mais, dans un certain nombre, plus faible de cas, c'est A_t qui sera supérieur à B_t et le jugement exprimé sera : A est plus favorable que B.

Si les deux distributions sont très éloignées l'une de l'autre, le jugement sera presque toujours le même. Plus elles se rapprocheront, plus la proportion des deux jugements ($A \succ B$ et $B \succ A$) tendra à s'égaliser. Si les moyennes des deux distributions (les « valeurs » des items) sont égales, on aura autant de jugements dans un sens que dans l'autre.

Le modèle permet donc de relier la distance entre les valeurs moyennes de A et de B et la probabilité que A soit jugé plus favorable que B. Si la distance est nulle, cette probabilité sera de O,5. Elle ne deviendra 1 que si cette distance est infinie. Réciproquement la connaissance de cette probabilité, estimée par la proportion de chacun des jugements, permet d'inférer la distance.

En pratique, comme on le fait très souvent en psychologie, on remplace l'analyse des fluctuations d'un seul individu par celle des différences entre les jugements d'individus différents, mais qu'on suppose homogènes. En d'autres termes, puisqu'il est pratiquement impossible de demander à une même personne de juger plusieurs fois, de façon indépendante, un ensemble d'items, on propose la même tâche à plusieurs individus dont on postule qu'ils sont identiques, les différences dans leurs jugements étant dues uniquement aux fluctuations qu'on a supposées chez un même individu. Ce postulat est très fort ; il admet en particulier que les différences entre les jugements de deux personnes sont purement aléatoires, n'ont donc aucune signification par rapport à l'attitude. En particulier, cela suppose que les jugements « objectifs » sont indépendants de la position personnelle de chaque juge, ce qui, empiriquement, ne se vérifie que dans des groupes par ailleurs déjà très homogènes. Il se peut que des groupes différents soient caractérisés par une organisation différente du champ étudié, par un système de valeurs différent. Par exemple, il existe des milieux, surtout dans les classes moyennes, ou chez les ouvriers, qui restent très attachés à un mode de vie traditionnel, mais qui sont en même temps influencés par les syndicats ou les partis de gauche ; pour eux, la libéralisation de la drogue, par exemple, sera jugée plus « à gauche » que certaines nationalisations. En revanche, dans des milieux dits « libéraux », on pourra tenir à la libre entreprise et en même temps tolérer des modes de vie très divers : les nationalisations leur apparaîtront comme plus révolutionnaires que la possibilité de fumer du haschisch. De telles différences, on le voit, sont d'une tout autre nature que les fluctuations aléatoires que postule le modèle.

En choisissant des juges dans ces diverses catégories, les fréquences de jugement ne feront peut-être que traduire l'importance de chacun dans l'échantillon des juges. Ou si, comme cela se fait souvent, le chercheur choisit, par commodité, les juges parmi ses collègues ou dans son entourage, il risque ensuite de classer l'ensemble de ses sujets, lors de l'analyse de l'enquête définitive, en fonction du seul système de valeurs de son propre milieu.

L'idéal consisterait à prendre séparément des groupes de juges dans les différentes catégories sociales visées par l'enquête, et à coder les réponses de chaque sujet en fonc-

tion des poids tirés des jugements des personnes de sa propre catégorie. Toutefois, outre que cette méthode est très lourde, elle rend impossible l'étude des attitudes de l'ensemble de la population selon une échelle commune à tous.

Nous retrouvons ici une difficulté que nous avons déjà signalée, et qui est liée à l'hétérogénéité fondamentale de la population, au moins pour de nombreux problèmes dont on traite par enquête, hétérogénéité qu'on ne peut pas toujours réduire à des différences quantitatives simples. Or, par définition, toutes les techniques de quantification, et en particulier les méthodes de construction des échelles d'attitude, ne s'appliquent qu'à de telles différences. Il ne faudrait toutefois pas en conclure que ces techniques sont, *en elles-mêmes*, entachées d'idéologie, comme on le voit parfois affirmé. Les méthodes sont ce qu'elles sont. En revanche, ce qu'on peut taxer d'idéologique, ou du moins ce qui demande réflexion, c'est l'application d'une technique à un domaine particulier, à la connaissance duquel on impose ainsi une certaine organisation, qui peut lui être tout à fait étrangère.

L'application de la méthode de Thurstone aux échelles d'attitude n'est pas la seule possible. Lorsqu'on estime que le postulat d'homogénéité de la population est satisfait, on peut utiliser cette méthode, ou du moins sa première phase, pour l'étude des préférences, des jugements esthétiques, pour constituer des échelles de prestige, etc. Elle n'est toutefois pas la seule possible, en particulier lorsqu'on dispose de comparaisons par paires. Différentes techniques ont été développées, les unes probabilistes, comme celles de Thurstone dont il est question ici, d'autres sans modèle probabiliste, vérifiant statistiquement l'hypothèse selon laquelle tous les jugements sont donnés au hasard, chacun avec la même probabilité que le jugement contraire. Le rejet de cette hypothèse amène à admettre que les jugements sont cohérents, à quelques erreurs sans signification près, et l'ordre sur les items est donné par celui du nombre de fois où chacun a été préféré aux autres. Cette méthode ne donnant pas de poids, de valeur à chaque item, mais seulement son ordre, ne peut être utilisée pour la seconde phase, pour distinguer les sujets eux-mêmes.

Pratiquement, la détermination du poids de chaque item à partir des jugements, comparatifs ou catégoriels, n'est pas difficile, du moins si l'on accepte quelques hypothèses simplificatrices. Revenons au modèle exposé ci-dessus. Il postule qu'un item sera jugé supérieur à un autre (ou placé dans une catégorie supérieure), si sa valeur instantanée est supérieure, ce qu'on peut écrire, si $A \succ B$ représente le jugement porté :

$$A \succ B \Leftrightarrow A_t > B_t \Leftrightarrow A_t - B_t = d_{AB} > 0$$

$$\text{Prob}(A \succ B) = \text{Prob}(d_{AB} > O)$$

Si les distributions des A_t et B_t sont gaussiennes, de moyennes V_A et V_B, et de variances σ_A^2 et σ_B^2, la distribution de σ_{AB} sera également gaussienne, de paramètres :

$\text{Moy}(d_{AB}) = V_A - V_B$

$\sigma^2(d_{AB}) = \sigma_A^2 + \sigma_B^2 - 2\gamma_{AB}\,\sigma_A\,\sigma_B$ si $_{AB}$ représente la corrélation possible entre les valeurs instantanées A_t et B_t.

Empiriquement, ce qu'on observe c'est la proportion de juges ayant donné $A \succ B$, donc une estimation de Prob ($A \succ B$).Si, pour simplifier, on admet que $\sigma^2(d_{AB}) = 1$, pour toute paire d'items, alors la figure 9 montre que, par simple consultation de la table de la loi normale réduite, on obtient une valeur $d_{AB} = V_A - V_B$.

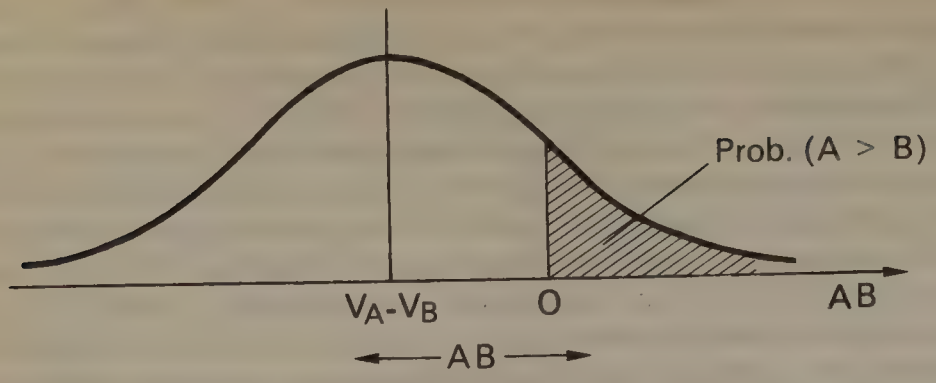


Fig. 9

On peut de cette manière calculer les distances entre toutes les valeurs moyennes des distributions des valeurs instantanées de chaque item.

En principe, si les jugements sont effectivement unidimensionnels, on doit observer, pour tout triple d'items A, B et C, la relation :

$(V_A - V_C) = (V_A - V_B) + (V_B - V_C)$

Même sans recourir à des tests plus précis, la simple inspection de ces relations peut donner des indications utiles et amener à rejeter le postulat d'unidimensionnalité.

Si ce critère est satisfait, au niveau d'approximation qu'on juge utile, on obtiendra les différentes valeurs V en posant que l'une d'elles est nulle. Si c'est V_A, on aura :

$V_A - V_B = - V_B$

$V_A - V_C = - V_C$ etc.

On a donc immédiatement une première approximation des valeurs de V_A, V_B, etc. Mais, connaissant V_B, nous pouvons tirer une nouvelle approximation de V_C de la valeur de $V_B - V_C$, et ainsi de suite. Si le modèle unidimensionnel s'applique, ces diverses estimations ne sont pas trop différentes et on pourra en prendre la moyenne.

Les valeurs V_A, V_B, V_C. correspondant à chaque item, constituent les poids recherchés. Remarquons que, alors que dans la méthode de Likert on attribuait un poids à chaque réponse, maintenant c'est à l'item lui-même qu'est associée une valeur. Au cours de la seconde phase, on attribuera à chaque sujet la somme des poids des opinions avec lesquelles il se sera déclaré d'accord, sans faire intervenir explicitement les autres [1].

d) L'analyse hiérarchique

Nous avons vu qu'on obtient une échelle hiérarchique (ou échelle de Guttman) avec deux questions dichotomiques lorsque, sur les quatre combinaisons de réponses possibles, trois seulement sont représentées, celle qui manque étant une des deux combinaisons intermédiaires. Cette définition peut se généraliser à toutes les paires présentant cette même relation entre elles. Dans ce cas, seules certaines combinaisons de réponses,

1. Cet exposé de la technique de Thurstone est particulièrement succinct. Nous recommandons au lecteur désireux d'aller au-delà du principe général et de l'application simple de se référer aux ouvrages cités dans la bibliographie de ce chapitre.

en petit nombre, sont représentées (n + 1 combinaisons, si on a n questions, sur les 2n possibles). Avec cinq questions, ce seront les six combinaisons suivantes :

Q_1	Q_2	Q_3	Q_4	Q_5	Rang
0	0	0	0	0	1
1	0	0	0	0	2
1	1	0	0	0	3
1	1	1	0	0	4
1	1	1	1	0	5
1	1	1	1	1	6

Si on admet que l'ordre 0 - 1 a la même signification pour toutes les questions, cet ensemble de combinaisons de réponses présente les propriétés suivantes :

— Les combinaisons sont totalement ordonnées, donc les sujets le sont aussi (plus exactement, ils sont pré-ordonnés).

— Les questions sont également ordonnées (si 1 représente l'accord avec une certaine opinion, on peut dire que Q1 est l'opinion la plus répandue, Q5 la moins répandue ; on pourra interpréter cela en disant que Q1 représente une opinion modérée, Q5 une opinion extrême ; s'il s'agit de problèmes à résoudre, l'ordre obtenu sera l'ordre de difficulté).

— Les sujets ayant donné la réponse 1 à la question de rang n ont également répondu 1 à toutes les questions précédentes.

— L'ordre obtenu sur les sujets est celui qu'on obtiendrait en associant aux réponses n'importe quel système additif de pondération.

— Les sujets classés au même rang ont tous donné exactement les mêmes réponses.

Le problème qui se pose maintenant est de savoir comment obtenir effectivement une échelle, ou, ce qui est un problème voisin, comment interpréter le fait que certaines questions constituent une échelle. Revenons pour cela au cas simple de deux questions seulement :

Q1	Q2
0	0
1	0
1	1

La réponse 1 à la question Q2 n'est donnée *que* par ceux qui ont aussi répondu 1 à Q1. On peut dire que, formellement, une des réponses *implique* l'autre. Les combinaisons de réponses présentes deviennent l'expression d'une relation logique entre les questions, la relation d'implication ; celle-ci étant une relation d'ordre, c'est elle qu'on interprétera ensuite comme un ordre sur les sujets ou sur les questions. Q1 et Q2 formeront une échelle si, logiquement, ceux qui ont répondu 1 à Q2 doivent avoir répondu également 1 à Q1, sans que l'inverse soit vrai.

Par exemple, on peut poser les deux questions suivantes :
Q1 : Achèterez-vous le produit S s'il coûte 10 francs ?
Q2 : Achèterez-vous le produit S s'il coûte 15 francs ?

Elle permettront, vraisemblablement, de distinguer trois catégories de personnes :

— Celles qui disent qu'elles n'achèteront pas S, quel que soit le prix (réponses « non » et « non »).

— Celles qui disent qu'elles achèteront S s'il coûte 10 francs, mais pas 15 (« oui » et « non »).

— Celles qui disent qu'elles achèteront S quel que soit son prix (« oui » et « oui »).

En revanche, il est peu probable qu'on trouve des personnes qui accepteraient de payer 15 francs mais pas 10. Payer 15 francs *implique* qu'on est également d'accord pour payer 10 francs.

Les trois catégories de personnes sont ordonnées selon l'attrait qu'exerce sur elles le produit en question : celles qui refusent de l'acheter même au prix le plus bas proposé peuvent être considérées comme celles pour lesquelles cet attrait est le plus faible ; pour celles qui l'achèteraient même au prix le plus élevé, cet attrait est le plus fort [1].

Il est toutefois assez rare que la relation entre deux ou plusieurs questions apparaisse comme aussi claire et contraignante : il arrive très souvent qu'on observe la même relation d'implication entre des questions non logiquement liées entre elles : l'existence d'une échelle traduit alors l'existence d'une cohérence plus générale ; inversement, dans certains cas, des relations logiques apparentes peuvent ne pas se vérifier, ce qui signale l'intervention d'autres facteurs. Dans notre exemple, si l'on trouvait une proportion non négligeable de personnes qui déclarent qu'elles achèteraient S à 15 francs mais pas à 10, cela pourrait indiquer que des éléments de prestige ou de consommation ostentatoire interviennent, ou qu'on ne croit pas qu'il s'agisse vraiment du même produit dans les deux cas, le prix plus élevé pouvant devenir un signe de qualité supérieure.

Le cas le plus fréquent, et le plus intéressant, est évidemment celui où l'existence d'une échelle hiérarchique traduit l'existence d'une cohérence socialement établie, mais peut-être momentanée, ou limitée à certains groupes seulement. Par exemple, actuellement, ceux qui s'affirment partisans d'une libéralisation de la législation sur la drogue se déclarent presque tous également anti-racistes, sans que l'inverse soit vrai : on a donc bien une échelle. Cette relation n'exprime aucune nécessité logique, mais correspond à une certaine constellation d'opinions et d'attitudes qui constituent l'idéologie de gauche d'aujourd'hui, qui est apparue récemment et pourra se modifier avec le temps.

Parlant plus haut du libellé des questions d'opinion (2.5.3.3.), nous avons distingué les opinions que nous avons appelées « *modérées* » des « *minimales* » [2]. Les questions minimales sont celles qui peuvent constituer une échelle hiérarchique. Le modèle sous-jacent, qui peut servir de guide dans le choix des questions à poser, est le suivant :

— Tous les sujets peuvent être ordonnés selon une variable, par exemple selon qu'ils sont plus ou moins favorables à un certain objet.

---

1. Cette interprétation de l'ordre suppose que la valeur de l'argent est la même pour tous les sujets interrogés, ce qui n'est pas nécessairement le cas. En toute rigueur, on ne saisit ici que les comparaisons entre le produit et les sommes d'argent indiquées. Si c'est la réaction au prix qui nous intéresse, ce peut être suffisant. Mais, si l'on veut saisir un attrait « absolu », en supposant que cela ait un sens, il faudrait faire intervenir d'autres dimensions que le prix.

2. Dans la littérature sur la question, on emploie en général plutôt « *questions ponctuelles* » et « *questions cumulatives* ». La terminologie que nous proposons ici nous semble présenter l'avantage d'attirer l'attention sur la signification de ces questions, et pas seulement sur leur forme.

— Les opinions qu'on peut énoncer sur cet objet peuvent elles aussi être ordonnées selon qu'elles lui sont plus ou moins favorables.

— Un sujet très favorable approuve toutes les opinions modérées, et d'autres, plus extrêmes ; il approuve donc toutes les opinions qu'approuvent aussi les sujets modérés.

— Un sujet peu favorable, par contre, n'approuvera que les opinions modérées et s'affirmera en désaccord avec les positions extrêmes.

Ce modèle suppose que *toutes les différences entre individus peuvent se ramener à des différences en plus ou en moins.* Il ne s'applique donc pas, par exemple, à une situation de polarisation, où certaines personnes acceptent tout ce que rejettent les autres, et il exclut aussi l'existence de positions intrinsèquement modérées, qui seraient qualitativement différentes des positions extrêmes, et pas seulement leur forme atténuée.

On voit que si l'on peut s'attendre à obtenir une échelle hiérarchique relativement à un objet bien déterminé ne suscitant pas de réactions trop fortes, le modèle devient difficilement applicable lorsqu'on veut étudier l'ensemble d'un champ complexe, comme par exemple celui des attitudes politiques. Non seulement il est peu probable que ce domaine soit unidimensionnel, mais il comprend des oppositions dont le schéma d'échelle, par sa nature même, ne peut pas rendre compte. Si l'on veut étudier, par exemple, la dimension gauche/droite, on ne pourra probablement pas la représenter par une échelle hiérarchique unique, il faudrait plutôt essayer de construire au moins deux échelles, l'une exprimant les opinions de gauche, l'autre les opinions de droite, et éventuellement une troisième pour le centre. Sur l'ensemble des opinions politiques, une échelle unique n'est pas possible, le schéma cumulatif ne pouvant pas s'appliquer.

En résumé, le modèle hiérarchique est plus apte à décrire une variation du type « moyennement favorable - très favorable », que la dimension « favorable - défavorable » ; et même dans le premier cas, il risque de négliger la spécificité des « moyennement favorables ».

Pratiquement, il est très rare qu'on puisse observer une échelle parfaite, correspondant en tous points au schéma que nous avons décrit. Il y a toujours un certain nombre de réponses non conformes, pour des raisons très diverses : existence d'autres formes de cohérence, incompréhensions, erreurs, etc. On est donc obligé de recourir à des approximations. Le principe général, sur lequel sont fondées toutes les techniques pratiques, consiste à rechercher, par tâtonnement ou par exploration systématique, un ordre des sujets et un ordre des questions qui donne une configuration la plus proche possible d'une échelle parfaite. On compte alors le nombre d'erreurs, c'est-à-dire le nombre de réponses qu'il faudrait modifier pour obtenir cette échelle parfaite. Si ce nombre est inférieur à 10 % du nombre total de réponses [1], on admet que c'est négligeable et qu'on

---

1. Cette valeur de 10 % est purement conventionnelle et ne constitue pas du tout un seuil de signification statistique. Elle est en général présentée sous la forme d'un « coefficient de reproductibilité » :

$$Cr = \frac{1 - \text{Nombre d'erreurs}}{\text{Nombre de questions} \times \text{Nombre de sujets}}$$

On admet que les réponses forment une échelle satisfaisante si la valeur de ce coefficient dépasse 0,9, ce qui revient exactement au même que de dire qu'il y a moins de 10 % d'erreurs.
D'autres critères, plus raffinés, ont été proposés. On les trouvera dans les ouvrages cités dans la bibliographie.

se trouve devant une approximation suffisante d'une échelle, et on effectue les « corrections » nécessaires.

Si le nombre d'erreurs dépasse ce seuil, on pourra rechercher si une partie seulement des questions ne présente pas une hiérarchie acceptable, ou si on ne trouverait pas une échelle parfaite dans une ou plusieurs sous-populations. Comme nous l'avons vu à propos de la méthode de Likert, il faut prendre garde à ces ajustements a posteriori, et toujours essayer de déterminer quelle peut en être la signification. Si, de toute façon, aucune échelle n'est acceptable, il faudra se demander si c'est parce que les questions choisies ne constituaient pas de bons indicateurs du domaine qu'on veut étudier, ou si c'est celui-ci qui n'est pas compatible avec la structure très particulière des échelles hiérarchiques [1].

Contrairement aux méthodes de pondération, qui acceptent les réponses telles que les sujets les ont données, l'analyse hiérarchique exige que l'ensemble des réponses des sujets présente une cohérence suffisante, les écarts par rapport à cette cohérence, les « erreurs », si elles sont en suffisamment petit nombre, étant purement et simplement annulées ; on admet ainsi qu'elles sont contingentes, sans aucune signification. Un exemple pris dans un domaine différent permettra de mieux saisir la différence entre les deux approches.

Supposons une série de problèmes à résoudre, de difficultés croissantes. Selon la méthode de Likert, on attribuera à chacun d'eux un poids d'autant plus élevé que moins de sujets l'auront résolu ; ce poids constituera donc une mesure de la difficulté. On pourra attribuer à chaque sujet une note égale à la somme des poids de tous les problèmes qu'il aura résolus. Si une personne a résolu un problème très difficile, mais a échoué à un autre, plus facile, elle aura une note plus faible que si elle avait résolu les deux.

Dans la perspective de l'analyse hiérarchique, en revanche, si les problèmes forment une échelle, cela signifie que ceux qui ont résolu un problème difficile auront presque tous aussi résolu les problèmes les plus faciles. Le sujet ci-dessus, qui a résolu un problème difficile et échoué à un problème facile sera considéré comme incohérent : on admettra que son échec devant le problème facile est attribuable à des facteurs contingents, qui n'ont rien à voir avec l'aptitude qu'on cherche à mesurer, et on le traitera comme s'il l'avait résolu. On privilégie donc la cohérence par rapport aux comportement effectifs.

Actuellement, la méthode de Likert est de loin la plus largement utilisée de toutes les techniques de construction d'échelles d'attitude. Elle ne demande pas de grands calculs, et ses hypothèses sont suffisamment faibles pour qu'elle soit compatible avec des données très diverses. L'analyse hiérarchique, en revanche, à cause de sa structure très particulière, jouit d'un grand prestige ; les échelles d'attitude construites de cette manière sont souvent jugées plus « solides » que les autres. En fait, nous l'avons vu, les exigences du modèle hiérarchique peuvent ne pas être adaptées au domaine étudié. Etre

1. Pour des détails sur la pratique de l'analyse hiérarchique, dont nous n'avons exposé que le principe général, on pourra se reporter aux ouvrages signalés dans la bibliographie à la fin de ce chapitre. Il existe actuellement des programmes d'ordinateur permettant de vérifier facilement l'existence d'échelles.

amené à rejeter l'hypothèse d'une échelle parfaite n'indique donc pas nécessairement que les questions choisies étaient mauvaises, mais simplement que la structure du domaine est différente, peut-être compatible avec des modèles eux aussi unidimensionnels, mais plus faibles.

Cette position est très différente de celle qu'on trouve souvent affirmée par les spécialistes, en particulier par Guttman lui-même. Pour lui, lorsqu'un ensemble de questions ne forme pas une échelle, il faut en conclure que le domaine qu'elles recouvrent n'est pas unidimensionnel, donc que l'attitude correspondante « n'existe » pas. A notre sens, cette conclusion est trop rigide : on peut toujours remettre en cause le choix des questions elles-mêmes, et aussi l'adéquation du modèle particulier, qui n'est pas la seule formalisation possible de la notion intuitive de dimension unique. Mais en même temps, comme nous l'avons vu, il faut se garder de trop d'ajustements ad hoc, qui risquent, pour des raisons purement formelles, de trop nous éloigner des hypothèses initiales.

e) Autres modes de construction d'indicateurs

Nous nous sommes un peu attardés, bien que nous en tenant à des généralités, sur les techniques de construction d'échelles d'attitude, parce que c'est dans ce domaine de la psychologie sociale que les méthodes ont été le plus développées et systématisées, ce qui a permis, à leur propos, de toucher à quelques problèmes plus généraux. Mais il ne faut croire ni que ces méthodes ne s'appliquent qu'à la mesure des attitudes, ni que ce sont les seules techniques possibles de construction d'indicateurs.

La structure d'échelle hiérarchique, par exemple, peut s'appliquer à bien d'autres matériaux que des opinions. On l'a appliquée à l'analyse de comportements politiques ou religieux, à des étapes de développement économique, ou, dans un domaine moins sociologique, à des stades de développement de l'intelligence. Quant aux méthodes de pondération, elles sont à la base de la grande majorité des indicateurs pratiquement utilisés, même si les méthodes de détermination des poids diffèrent de celles que nous avons vues.

En fait, même lorsqu'on n'applique pas rigoureusement les méthodes que nous avons exposées, la construction d'indicateurs recourt presque toujours à deux opérations fondamentales, qu'on peut rapprocher de ces méthodes :

a) *La composition complète de tous les indicateurs élémentaires.* Cela se fait le plus souvent par addition, mais on peut aussi prendre le carré des valeurs recueillies, ou leur logarithme. Quel que soit le mode de composition adopté, il faut attribuer un poids à chacun des indicateurs élémentaires ; ne pas le faire revient à attribuer à chacun le même poids, ce qui est un système de pondération comme un autre. Comme nous l'avons vu à propos des méthodes de pondération, cette composition ne va pas nécessairement de soi : elle demande à être justifiée, pour éviter de combiner des éléments qui, en fait, n'ont rien de commun, et dont la composition n'aurait aucune signification. Cette justification peut se faire soit empiriquement, en montrant que les indicateurs élémentaires sont suffisamment en corrélation les uns avec les autres pour qu'on admette qu'ils saisissent bien une même réalité sous-jacente, soit en s'appuyant sur une théorie préalable.

b) *La prise en considération de l'indicateur élémentaire le plus significatif.* On caractérise chaque unité par l'indicateur élémentaire ayant la valeur la plus forte. Par exemple, si on veut construire un indicateur d'intensité de la pratique religieuse à partir de questions sur différents comportements, on pourra attribuer à chaque sujet un niveau correspondant à sa pratique la plus « forte ». Si quelqu'un va à la messe chaque dimanche, on le considérera comme plus pratiquant que celui qui ne fréquente l'église qu'aux grandes occasions. Cette manière de faire admet implicitement l'existence d'une échelle hiérarchique : normalement celui qui va à l'église chaque dimanche fait également baptiser ses enfants. Mais, souvent, les comportements effectifs ne présentent pas une telle cohérence, et ne constituent donc pas une échelle. Mais on conserve la référence implicite à une telle structure hiérarchique, en négligeant les « incohérences », même si elles sont nombreuses. Mais cela suppose qu'on dispose d'un critère externe d'importance des comportements.

On voit que le chercheur dispose ici d'une très grande latitude, qui peut vite devenir dangereuse. A titre de garde-fous, il faudrait toujours, lorsqu'on construit ainsi un indicateur à partir de jugements plus ou moins informels et intuitifs, se poser des questions telles que :

— Est-ce que les sujets que je confonds (que je situe dans la même catégorie) ont vraiment suffisamment de traits importants en commun ? Quelle information est-ce que je perds en ne les distinguant pas ?

— Les relations entre les groupes ainsi constitués correspondent-elles à la représentation que je me fais du phénomène étudié ? sinon, pourquoi ?

— Les valeurs extrêmes de l'indicateur correspondent-elles bien aux valeurs extrêmes du phénomène qu'on veut décrire ?

— D'autres indicateurs sont-ils concevables ? Quels classements donnent-ils ? Quelles sont les relations entre ces divers indicateurs ? Quels sont les sujets qui ne sont pas classés de la même manière par ces divers indicateurs ? Pourquoi ?

Les techniques de construction d'échelles d'attitude visent à obtenir, à partir des indicateurs élémentaires, immédiatement observables, des indicateurs complexes constituant des catégories ordonnées. Dans certains cas, on peut très bien ne pas s'imposer cette contrainte. Par exemple, si l'on veut combiner en un indicateur unique deux indicateurs élémentaires ordinaux, on pourra simplement les croiser, et regrouper, selon des hypothèses extérieures, ou la considération des fréquences dans chaque cas, différents éléments du produit cartésien ainsi obtenu, ce qui donnera un ordre partiel sur les catégories constituées, mais on pourrait les traiter simplement comme des catégories qualitativement différentes.

7.2.2. Les variables dites classiques

7.2.2.1. Intérêt et limites de l'emploi de ces variables

Les méthodes que nous avons exposées au cours du chapitre précédent permettent de construire des variables ad hoc, indicateurs des phénomènes qu'on veut décrire et

analyser, ou de leurs facteurs explicatifs. Mais il est extrêmement rare qu'au cours d'une enquête, surtout s'il s'agit d'un travail « appliqué », on construise de cette manière toutes les variables qui interviendront dans les analyses. Tout questionnaire, ou presque, comporte, traditionnellement, un certain nombre de questions dites d'*identification*, portant sur le sexe, l'âge, la catégorie sociale, le revenu, le statut matrimonial, la taille de l'agglomération habitée, parfois aussi le niveau d'instruction ou le nombre d'enfants. Aux Etats-Unis, on ajoute presque toujours à cette liste la race, et souvent la religion.

Il est très fréquent que l'analyse des résultats d'une enquête se limite aux croisements entre ces variables générales et les questions spécifiques au problème étudié. Une étude d'opinion publique, par exemple, consiste souvent dans l'étude, et éventuellement l'interprétation, des relations entre les réponses à des questions portant sur un problème actuel et ces variables d'identification, auxquelles on pourra ajouter un indicateur de la position politique. De même, une étude de marché se réduit parfois à la mise en évidence de relations entre ces mêmes variables et certaines consommations, préférences ou intentions d'achat. Même lorsque l'ambition du chercheur est plus grande, et que l'exploitation des résultats est plus sophistiquée, ces croisements avec les variables « classiques » sont pratiquement toujours effectués et analysés.

Cette pratique est tellement générale qu'elle semble aller de soi. On en rencontre néanmoins quelques critiques, mais elles sont en général incluses dans une critique globale de l'enquête en tant que méthode. Pour certains, on n'a rien expliqué quand on s'est contenté d'affirmer qu'une personne agit d'une certaine manière « parce qu'elle a entre 25 et 34 ans » ou « parce que c'est un ouvrier », et encore moins quand on emploie le langage de la causalité, et dit que la catégorie sociale ou l'âge sont des « causes » de comportement. De ce point de vue, l'explication qui seule a un sens, c'est celle qui dit qu'un individu agit ainsi parce qu'il est dans une certaine situation ayant pour lui une signification déterminée, et parce qu'il poursuit un certain but, en se conformant à certaines normes. Le comportement a des *raisons* non des *causes*.

Ce type d'objection renvoie à une opposition théorique profonde, qui porte sur la conception même de l'explication dans les sciences sociales, et à la limite sur l'objet même de la sociologie ; il dépasse donc de très loin le problème des enquêtes. Il s'agit d'un point de vue très proche de celui de Max Weber et de sa sociologie compréhensive, et de celui de Wittgenstein, pour qui le sens des conduites ne peut s'appréhender que du point de vue du sujet agissant [1].

C'est évidemment en opposition à la fois avec le point de vue durkheimien et avec le point de vue marxiste, qui insistent au contraire sur l'impossibilité pour un individu de prendre vraiment conscience des déterminismes auxquels il est soumis. Mais on peut se demander si l'enquête, en tant que méthode de questionnement, est liée de façon nécessaire à une théorie sociologique particulière, ou si ce n'est pas dans le choix des hypothèses, des variables explicatives et surtout, plus en amont, dans la manière de poser le problème, que le poids de la théorie intervient. Rien n'empêche, en principe, d'introduire dans une enquête des questions portant sur les représentations et les inten-

1. On pourrait aussi rapprocher de ce point de vue les idées de Popper sur la « Logique de la situation ».

tions. Toutefois, en fait, la pratique des enquêtes, bien que souvent présentée comme purement empirique, et donc théoriquement neutre, est en général, par ses présupposés implicites, proche d'explications de type durkheimien, considérablement simplifiées.

Le choix de faire jouer un rôle central aux variables d'identification, que le jargon usuel appelle souvent « classiques » ou « objectives » ce qui est significatif, relève à la fois d'une sociologie implicite, et d'un certain nombre de contraintes sociales liées au fonctionnement de la communauté des chercheurs spécialisés, à une « forme de vie » (*form of life*), aurait dit Wittgenstein.

Le recours privilégié à un petit nombre de variables considérées comme explicatives, toujours les mêmes, renvoie à une conception implicite de ce qu'« est » un individu : c'est quelqu'un qui est d'un certain sexe, a un certain âge, a une certaine activité professionnelle, est marié ou non. On travaille avec ces variables comme si elles définissaient une personne, et surtout la définissaient « objectivement ».

Ces variables, en effet, apparaissent comme des données objectives qui *sont là* et ne posent aucun problème de définition ou d'interprétation. A force de les utiliser, on oublie par exemple que le code des catégories socio-professionnelles utilisé en France est de création récente, et propre à ce pays, qu'une variable aussi simple que l'âge peut être traitée de plusieurs manières différentes, et que même la définition du sexe peut être l'objet de discussions. Loin d'être des données objectives, ces variables sont l'aboutissement de processus sociaux de définition. Mais la caractéristique de ces processus est précisément de les objectiver, ce qui en fait une réalité en soi, sur laquelle d'ailleurs, dans notre société, le consensus peut être quasi général, partagé à la fois par le sociologue, ses lecteurs et ses sujets.

Le processus d'objectivation est complexe, et mériterait d'être étudié pour lui-même, mais on peut se rendre compte de son existence par des détails très simples, et qui passent souvent inaperçus. Par exemple, celui qui voudrait utiliser, en France, une classification socio-professionnelle différente de celle de l'I.N.S.E.E., parce que celle-là lui semblerait mieux adaptée à son problème, se heurterait à des obstacles multiples. Tout d'abord, faute d'informations préalables sur ces nouvelles catégories, il ne pourrait pas stratifier son échantillon, ni indiquer de quota à ses enquêteurs. Une fois l'échantillon constitué, il ne disposera d'aucune donnée à laquelle le comparer pour vérifier s'il ne présente pas de biais imprévu. Dans l'analyse des résultats, il ne pourra pas s'appuyer sur des travaux antérieurs, comparer ses conclusions à celles d'autres chercheurs. De plus, utilisant ces catégories pour la première fois, il n'en aura aucune expérience, alors que les variables classiques évoqueront pour lui de nombreux autres travaux, portant sur le thème de sa recherche ou sur d'autres. S'il se pose des problèmes de prévision, il ne trouvera aucune série chronologique indiquant comment la répartition de la population entre ces nouvelles catégories a évolué dans le passé, ce qui pourrait lui fournir des hypothèses sur leur évolution future. Supposons qu'il soit publicitaire : son enquête devrait l'amener à définir une « cible » en fonction d'habitudes de consommation, de préférences, etc. Mais les données sur la couverture des différents supports publicitaires donnent la clientèle de ceux-ci, analysée selon les variables classiques, dont la C.S.P., et évidemment pas selon des variables plus spécifiques, mais pertinentes pour un seul problème. Pour élaborer un plan de répartition de la publicité, on ne peut pas

éviter de définir la cible en fonction de variables classiques quitte, s'il le faut, à la simplifier en regroupant plusieurs catégories, ou à la raffiner en décomposant certaines.

Socialement, le code des C.S.P. s'impose donc au chercheur. S'il n'en est pas satisfait, il peut le compléter par d'autres variables mieux adaptées au problème précis qu'il traite ou à sa position théorique, mais il s'en passera difficilement, sauf à avoir une problématique tellement personnelle que l'isolement conceptuel ne le gênera pas, bien qu'il puisse constituer un obstacle à la bonne compréhension par le lecteur.

Pourtant, le code des C.S.P. n'a rien d'absolu, ni même de théoriquement fondé ; il est propre à la France, où il a été élaboré à la fin des années quarante, surtout par des statisticiens et des économistes, et depuis il a fréquemment été critiqué. Aux Etats-Unis, par exemple, on utilise habituellement une échelle de « niveaux de statut », en six catégories ordonnées. Mais, malgré les critiques qui lui ont été adressées, et le fait que, la structure socio-professionnelle ayant changé en trente ans, les catégories ont vieilli, on continue à l'utiliser. Le souci d'accumulation de résultats comparables, la possibilité de suivre des évolutions au cours du temps, ont jusqu'à présent prévalu contre toutes les propositions de modification, même quand on les reconnaissait fondées. Et si, comme cela arrivera probablement, un nouveau système, tenant compte à la fois de l'évolution de la société et de celle des perspectives théoriques, est élaboré, il jouera le même rôle contraignant et unificateur que le système actuel. Vraisemblablement d'ailleurs, il faudra pendant un certain temps utiliser simultanément les deux systèmes, afin de bien étudier leurs correspondances et assurer la continuité des résultats ainsi produits.

Le fait que ces variables sont construites, puis traitées comme « objectives », c'est-à-dire détachées du processus de leur construction, est moins visible pour d'autres variables, apparemment plus simples, comme l'âge ou le sexe. Le consensus sur leur définition est tel que, même si des problèmes se posent devant des cas très marginaux, ceux-ci restent extrêmement rares. Toutefois, si l'attribution d'un individu à une classe d'âge ou de sexe peut se faire facilement et sans ambiguïté, la signification de ces variables est loin d'être univoque.

Les différences entre les hommes et les femmes ne se limitent pas à quelques organes ou à des hormones. Ce sont aussi des différences de rôle et de statut, qui elles-mêmes varient selon les catégories sociales ou l'âge, des différences d'activités professionnelles, de revenus, etc. Interpréter une différence observée entre un échantillon d'hommes et un échantillon de femmes, c'est faire appel à tout ce qu'on sait à ce propos, c'est-à-dire, en définitive, à toutes les connotations qu'évoquent pour le chercheur les mots « homme » et « femme », ces connotations provenant à la fois de l'expérience quotidienne du chercheur et de son expérience professionnelle. Lorsqu'on ne dispose pas de telles connotations ni d'une théorie explicative, la seule constatation d'une telle différence n'apporte rien. C'est le cas, par exemple, du *rang de naissance*, qu'on a constaté être en corrélation avec de nombreux comportements ou attitudes, aussi divers que le nombre d'accidents de voiture ou la tendance à manipuler les autres. Mais, jusqu'à présent, quelles qu'aient été les hypothèses initiales, on ne peut pas dire que ces corrélations aient contribué à éclairer ces problèmes, le rang de naissance n'évoquant pas suffisamment d'autres caractéristiques pour qu'on puisse « expliquer » ces rela-

tions, et aucune théorie convaincante n'ayant été proposée à ce sujet. Toutefois, si l'on continue à introduire cette variable dans les enquêtes, et à accumuler des observations à son propos, elle pourra finir par acquérir suffisamment de connotations pour devenir explicative, et éventuellement pour qu'on puisse l'intégrer dans une théorie explicite et plus générale.

On voit que le caractère « objectif » de la mesure d'une variable, c'est-à-dire l'absence d'ambiguïté du processus d'attribution d'un individu à une catégorie de cette variable, n'implique ni la neutralité théorique de la constitution de ces catégories, ni l'objectivité de l'utilisation qui en sera faite. Les variables « classiques » ont acquis leur objectivité d'abord à cause de la facilité de la mesure, et aussi parce que ce sont celles qui servent à catégoriser les individus dans la vie quotidienne ; puis la large utilisation qui en a été faite par les chercheurs a contribué à cette objectivation, cette utilisation venant renforcer l'expérience quotidienne par la connaissance de résultats plus scientifiques.

Certes, si l'on disposait de variables, de catégories mieux fondées sur une théorie, le problème se poserait autrement. Mais même dans les théories où ce problème est essentiel, comme le marxisme où les classes jouent un rôle central, la définition empirique des catégories reste insuffisante, et la correspondance entre les notions théoriques et les observables reste incertaine. La situation que nous venons de décrire, avec tout ce qu'elle peut avoir d'insatisfaisant, est le reflet des insuffisances théoriques actuelles de la sociologie.

De ces remarques, il ne faut toutefois conclure ni que l'usage de variables classiques est totalement arbitraire et injustifié, ni qu'il est absolument contraignant. Il est peu probable, par exemple, qu'une revue scientifique refuse la publication d'un compte rendu d'enquête qui n'utiliserait pas les variables en question ; simplement, l'utilisation par d'autres chercheurs des résultats obtenus risquera d'être plus limitée. D'ailleurs, comme nous l'avons dit, rien n'empêche le chercheur qui le juge nécessaire d'ajouter d'autres variables plus conformes à ses objectifs propres, ni d'effectuer des regroupements ou des décompositions des catégories traditionnelles.

Peut-être, paradoxalement, la principale justification de l'emploi généralisé des variables classiques, en plus de son effet unificateur, réside précisément dans le fait qu'il ne s'agit pas de variables « pures », construites en fonction d'une théorie. Cela leur permet de fonctionner comme des *indicateurs* de situations sociales diverses ; interpréter une relation avec l'une de ces variables, lorsqu'on se donne la peine de le faire et qu'on ne se borne pas à constater l'existence de la relation, revient presque toujours à faire des hypothèses sur l'aspect de ces situations qui a des chances d'être pertinent pour le problème traité. En effet, sociologiquement et psychologiquement, ces variables sont composites : selon le cas, une différence observée entre hommes et femmes pourra être interprétée comme une différence dans la situation économique ou professionnelle, comme l'expression de leurs rôles sociaux respectifs, comme la conséquence de la diversité de leurs contraintes de temps, etc. De même, être « jeune », c'est à la fois être dans un certain état physique, être étudiant ou au début de sa vie professionnelle, avoir un salaire faible, être célibataire, participer à une sous-culture différente de celle de ses aînés, etc. On voit que le « sexe » ou « l'âge » en soi, ne signifient rien. Ou

plutôt, ils signifient trop de choses diverses pour qu'on puisse se dispenser d'approfondir et de spécifier la relation entre une de ces variables et un comportement ou une opinion, ce qui revient à préciser laquelle des *composantes* de cette variable intervient. Sans une interprétation de ce type, la corrélation n'aura aucune valeur explicative, aucune portée théorique, et parfois aucune utilité pratique non plus.

Mais, en général, cette analyse en termes de composantes n'est pas faite explicitement. Ce sont plutôt les représentations du sociologue, représentations qui sont d'ailleurs fondées autant sur une expérience quotidienne que sur des connaissances scientifiques, qui le guident dans ses interprétations. Si on veut dépasser le stade des hypothèses, il faut disposer de mesures indépendantes de la composante en question, et mettre statistiquement en évidence son influence (cf. 7.3.1.1.).

Envisagées de cette manière, les variables classiques peuvent être effectivement considérées comme de bons indicateurs de la situation dans laquelle se trouvent les individus, et la relative simplicité de leur définition et de leur mesure devient un avantage non négligeable, de même que le caractère cumulatif des enquêtes que permet leur emploi généralisé.

7.2.2.2. Les principales variables classiques

a) La triade sexe - âge - niveau

Si les variables dont il est question ici ont acquis, en partie grâce à leur emploi répété, le statut d'indicateurs, on peut aussi donner à l'usage de certaines d'entre elles un fondement théorique plus solide, en montrant qu'elles constituent effectivement des facteurs de différenciation essentiels. G. Balandier considère comme fondamentales, dans toute société, les différences d'*âge*, de *sexe* et de *position sociale*, et reconnaît une importance aussi grande aux oppositions entre « classes sexuelles » et entre « classes d'âges », qu'aux conflits entre catégories sociales. De même, extrapolant à partir des études récentes sur les sociétés de primates, E. Morin distingue chez elles trois groupes : les mâles adultes qui dominent les autres, les femelles et les jeunes mâles, avec en plus, une relation hiérarchique qui semble exister à l'intérieur du groupe des mâles adultes, avec des conséquences sur les relations à l'intérieur des autres groupes.

Si les marxistes, évidemment, attribuent une importance essentielle aux oppositions entre classes sociales, ils tendent de plus en plus à prendre en considération les deux autres, même si c'est pour montrer qu'elles sont, en dernière analyse, des expressions de la lutte des classes, fondamentale [1].

Il y a donc une certaine convergence, chez les sociologues, pour reconnaître l'importance de ces trois dimensions, qu'on parle à leur propos d'oppositions, de conflits, ou simplement de différences. Prendre comme facteurs de base d'une analyse sociologique les variables classiques, qu'on peut y réduire approximativement, n'est donc pas injustifié, à condition qu'on ne s'y borne pas.

Toutefois, le type d'analyse le plus couramment effectué peut amener à masquer les

1. On peut se demander si le fait que les sociologues, marxistes ou autres, accordent maintenant une importance *théorique* au sexe et à l'âge n'est pas simplement le reflet des mouvements sociaux actuels.

phénomènes d'opposition liés à ces facteurs. En effet, on a tendance à traiter la plupart des variables comme des variables continues, les différences entre catégories successives, par exemple entre catégories d'âge ou de revenu, étant considérées comme simplement quantitatives. Or, cela rend impossible la mise en évidence de groupes ayant des situations, des représentations, des attitudes, des comportements, qualitativement distincts, c'est-à-dire ne pouvant pas être simplement décrits par des différences en plus ou en moins. Pour pouvoir étudier des groupes sociaux, et pas simplement la variable « niveau social » par exemple, il faut partir d'hypothèses relatives à ces groupes, et les étudier, s'il le faut, séparément, en retenant pour chacun les traits qui leur sont pertinents. Rien ici n'est incompatible avec les techniques d'enquête, mais cela exige des analyses complexes.

Une autre difficulté de l'analyse au moyen des variables classiques provient de ce qu'elles sont en corrélation les unes avec les autres, ce dont on omet parfois de tenir compte. Même théoriquement, dans la conception de Balandier, par exemple, les trois dimensions qu'il distingue ne sont pas indépendantes : les hommes et les femmes, les jeunes et les vieux, n'ont pas le même pouvoir, n'occupent pas la même situation sociale. Empiriquement, on le retrouve dans les données sur la population des pays développés. On sait, par exemple, que les revenus personnels des hommes et des femmes sont très différents, et que ce même revenu varie en fonction de l'âge. L'extension récente des études secondaires et supérieures fait apparaître, dans l'ensemble de la population, une corrélation inverse entre l'âge et le niveau d'instruction, etc.

Il en résulte que l'interprétation directe d'une corrélation simple entre une réponse quelconque et une seule de ces variables, est à peu près impossible, en toute rigueur. Il faut se donner les moyens de séparer, statistiquement, les effets de ces différents facteurs, ce qui exige des échantillons importants et des analyses qui, sans être très sophistiquées, sont un peu plus complexes que le simple croisement de deux variables. En ne prenant pas ces précautions, on risque de se fourvoyer complètement dans les interprétations, attribuant, par exemple, à l'âge ce qui est dû en réalité au revenu ou au niveau d'instruction.

b) Le niveau social

La détermination des catégories de niveau constitue un des problèmes centraux de la sociologie. Qu'on parle de hiérarchie, de stratification, de classes, il s'agit toujours de systématiser l'idée que les individus peuvent être regroupés en catégories qui diffèrent par leurs ressources, leur pouvoir, leur statut, leur culture, etc., et que cette appartenance détermine un grand nombre de comportements et d'attitudes. Il n'est pas dans notre propos, ici, d'aborder le problème théorique des inégalités et de la différenciation sociale, ni même de passer en revue les différentes solutions qui ont été proposées.

Dans la pratique des enquêtes, ces inégalités sont le plus souvent saisies par trois variables fortement corrélées : le revenu, la Catégorie Socio-Professionnelle (C.S.P.) et le niveau d'instruction. Pour désigner ce qu'on cherche à saisir ainsi, on emploie souvent les termes « niveau socio-économique » ou « niveau socio-culturel », ce qui exprime bien que, de façon relativement simple, on essaie de saisir une réalité complexe à plusieurs dimensions, qu'on tente de réduire à une seule.

Le *revenu* est, à première vue, la variable la plus simple et la mieux définie. Il a une signification claire et évidente de mesure des ressources financières ou, réciproquement, de contrainte ; mais, de plus, il peut être interprété comme indicateur de niveau social et de mode de vie. Malgré sa simplicité, la définition de revenu n'est pas unique, et exige d'être précisée.

On peut prendre en considération le *revenu personnel* de la personne interrogée, ou le revenu *total du ménage* auquel elle appartient. Si l'on s'intéresse en premier lieu à l'aspect économique des comportements, et qu'on veut une estimation des ressources, c'est évidemment le revenu total du ménage qui doit être pris en considération, en n'oubliant pas d'y inclure les prestations familiales, les primes régulières ou occasionnelles, les revenus annexes, etc. Mais pour une évaluation correcte des disponibilités, des possibilités de dépense ou d'épargne effectives pour ce revenu, il faut aussi tenir compte de la taille du ménage. A revenu égal, un célibataire n'ayant personne à sa charge, et un ménage de plusieurs personnes, n'ont évidemment pas les mêmes disponibilités. Pour corriger cette erreur, on calcule parfois le *revenu moyen par tête*, ou, ce qui est plus raffiné, le *revenu par unité de consommation* qui est une moyenne pondérée, où tous les membres du ménage n'ont pas le même poids (par exemple, dans un des systèmes d'unités de consommation proposés, le premier adulte compte pour 1, les suivants pour 0,8, les enfants de moins de 15 ans pour 0,5).

Lorsque, plus ou moins clairement, l'interprétation dépasse la seule signification économique du revenu, lorsqu'on le prend comme indicateur des variables plus sociologiques, considérer le revenu total du ménage peut masquer des distinctions importantes : on confondra, comme ayant le même revenu global, un ménage où une seule personne travaille et touche un revenu élevé, et un autre où la même somme est obtenue par l'addition de deux, ou parfois trois, revenus faibles. Il est évident que les deux ménages occupent des positions sociales différentes, et on peut s'attendre que, sur la plupart des points, les comportements de leurs membres soient différents.

Pour certains problèmes, en particulier tous ceux qui ont trait à la vie professionnelle, c'est le revenu de la personne interrogée elle-même qu'il faut prendre en considération. Dans d'autres cas, il pourra y avoir intérêt à connaître également le revenu du chef de ménage. Par exemple, si une femme mariée se contente d'un salaire d'appoint assez faible, son mari ayant un revenu élevé, c'est celui-ci qui permettra le mieux de la situer socialement, plutôt que le sien propre.

Lorsqu'on demande à une personne d'indiquer son revenu, il est important, pour éviter des équivoques, de lui indiquer ce qu'on entend par-là. Est-ce uniquement le produit régulier de son travail, ses autres revenus éventuels (pensions, revenus immobiliers, etc.), la totalité de ce qu'elle perçoit au cours d'une année ? Selon le cas, on s'intéressera à l'une ou l'autre de ces valeurs, mais il est important de toujours connaître exactement la signification du nombre obtenu. Pour éviter les malentendus, on peut présenter à l'enquêté une liste de tous les éléments de revenus, dont on lui demande de tenir compte. Mais, même avec cette aide, la somme à laquelle on arrive peut être très erronée. Si, pour les salariés, il est relativement facile d'estimer les différentes composantes du revenu (salaire de base, heures supplémentaires, primes, prestations fami-

liales), il n'en va pas de même des différentes catégories de travailleurs indépendants ou de patrons, dont les revenus sont irréguliers et mal connus, même par eux.

Mais la principale difficulté dans l'estimation du revenu comme variable explicative, c'est la réticence d'une part importante des Français. Dans une enquête portant sur l'ensemble de la population française, on rencontre environ 20 % de personnes qui refusent d'indiquer leur revenu. C'est la somme de refus de types différents : ceux qui estiment que c'est une question indiscrète, et qu'on trouve dans presque toutes les catégories, ceux qui ont des revenus soit très bas, soit très élevés, et qu'une forme de pudeur empêche d'avouer, les indépendants enfin, qui craignent une indiscrétion en direction du fisc. Certaines femmes mariées ignorent ce que gagne leur mari, et ne peuvent donc pas donner d'indications sur le revenu du ménage.

Parmi ceux qui donnent une réponse, il n'est pas certain que les estimations soient correctes ; les commerçants, par exemple, indiquent souvent leur forfait fiscal, qui peut être différent de leur revenu réel. Tout ceux qui, d'une manière ou d'une autre, essaient de frauder le fisc, donnent une valeur plus proche de leur déclaration que de la réalité. Enfin, on peut supposer que des personnes craignent, pour des raisons diverses, d'apparaître comme ou trop pauvres, ou trop riches, et biaisent leurs réponses en conséquence. On connaît les difficultés qu'a l'administration à connaître les revenus réels de la population ; on retrouve les mêmes problèmes dans les enquêtes, la crainte d'une indiscrétion restant très forte malgré les assurances de secret professionnel qu'on peut donner.

La question sur les revenus étant souvent jugée très indiscrète, on la place, habituellement, à la fin du questionnaire, pour éviter de bloquer trop tôt le sujet. On peut aussi présenter la question sous une forme fermée, en présentant une fiche pré-établie avec une liste d'intervalles de revenus, en demandant simplement à la personne interrogée d'indiquer dans quelle catégorie elle se situe, et non son revenu exact, ce qui semble mieux accepté qu'une simple question directe.

La saisie directe de la *catégorie sociale* pose des problèmes conceptuels complexes, et les sociologues sont loin d'être d'accord entre eux sur la définition. En France, la classification la plus largement utilisée est celle des Catégories Socio-Professionnelles (C.S.P.), élaborée par l'I.N.S.E.E. Sous sa forme la plus simple, elle comprend dix catégories :

— Salariés agricoles
— Agriculteurs exploitants
— Commerçants, artisans
— Patrons de l'industrie et du commerce
— Cadres supérieurs, professions libérales
— Employés
— Cadres moyens
— Ouvriers
— Divers (personnel de service, armée, etc.)
— Inactifs, retraités.

On voit que c'est une classification fondée sur l'*activité professionnelle*, définie à la fois par sa nature et par son niveau. On peut donc la rapprocher de la notion marxiste

de classe, définie à partir de la place dans le système de production. Mais on y trouve aussi, sous-jacente, l'idée d'une différenciation fondée sur le revenu et sur le niveau culturel, ainsi que la distinction entre les secteurs primaire, secondaire et tertiaire. Des classifications plus fines, construites par subdivisions de celles-ci, sont également disponibles, bien que moins souvent utilisées. En principe, ces catégories ne sont pas ordonnées ; en fait, dans les interprétations, ou même dans certains traitements, on se réfère souvent, implicitement, à un ordre qui correspond à la fois à celui des revenus et à une représentation de la hiérarchie sociale.

Cette classification a souvent été critiquée. Par exemple, on a fait remarquer que considérer les « employés » comme constituant une catégorie amène à regrouper des personnes aussi diverses que certains techniciens et des vendeurs de grands magasins. Constituer une catégorie unique avec les cadres et les professions libérales est justifié du point de vue du niveau culturel et de certains aspects du mode de vie, mais pas du tout si on s'intéresse à leur rôle socio-politique ou économique. Il faut toutefois noter que si on a besoin de ces distinctions, on peut toujours recourir aux classifications plus fines.

Définir la position sociale d'une personne ou, ce qui est un autre point de vue sur le même phénomène, les groupes sociaux importants, est un problème sociologique fondamental, certains disent même que c'est le problème de base de la sociologie, mais c'est aussi un des points sur lequel le consensus est loin d'être atteint. La définition des classes par Marx est difficile à utiliser empiriquement, et les controverses, même entre marxistes, ne manquent pas sur ce point. La définition à partir de la seule place dans le système de production, ou ses dérivés, est d'ailleurs souvent remise en cause. Définir des catégories par le mode de consommation, par le caractère plus ou moins moderne de l'activité professionnelle, par la proximité par rapport à divers mouvements sociaux, a été proposé, et a pu susciter des travaux intéressants, mais sans qu'aucune de ces classifications ne se soit imposée.

Le *niveau d'instruction*, appelé aussi parfois abusivement *niveau culturel*, est très corrélé aux deux variables précédentes, et constitue, au-delà de sa signification précise, un bon indicateur du « niveau » que l'on cherche à saisir.

On l'évalue soit sous la forme de l'âge de fin d'études, soit du diplôme le plus élevé atteint, ou encore du niveau du dernier établissement fréquenté. Si l'on se contente de trois grandes catégories, niveaux primaire, secondaire et supérieur, on dispose d'une variable ordinale. Si l'on essaie de raffiner la mesure, et de tenir compte de la nature des études suivies et des diplômes obtenus, il devient beaucoup plus difficile de définir un ordre. En particulier ceux qui ont fait des études techniques au-delà du C.A.P. sont très difficiles à situer sur une échelle.

On dispose de beaucoup moins de données sur le niveau d'instruction que sur la catégorie socio-professionnelle, ce qui peut parfois rendre son emploi difficile. En revanche, il semble constituer un bon indicateur de milieu et de mode de vie.

Nous avons vu, à propos du revenu, qu'on peut hésiter entre le revenu personnel de la personne enquêtée ou le revenu total du ménage. Des problèmes du même type se posent pour la C.S.P. et le niveau d'instruction. Les notions correspondantes n'existent évidemment pas pour la famille globalement, mais on peut se demander si, pour certains problèmes, celle-ci, ou ses différents membres, ne sont pas mieux caractérisés par

la C.S.P. ou le niveau d'instruction du chef de ménage. Dans l'incertitude, il est préférable de recueillir les deux informations, et de voir après coup laquelle présente la plus forte corrélation avec les traits que l'on veut expliquer. Constater que c'est l'une ou l'autre de ces variables qui est la plus « explicative » peut constituer en soi une indication intéressante.

Dans certains cas, les caractéristiques des parents peuvent constituer un indicateur utile du niveau culturel, même lorsqu'il s'agit de personnes adultes. Par exemple, on a pu montrer que certains aspects des goûts esthétiques sont mieux corrélés avec le niveau d'instruction du père qu'avec celui de la personne interrogée elle-même, ce qui s'explique si l'on pense que ces goûts se forment probablement très tôt. La différence entre le niveau socio-économique ou socio-culturel des parents et celui de la personne interrogée peut aussi caractériser utilement celle-ci, en la situant dans une dynamique sociale.

c) L'âge et les générations

Bien que ce soit une des variables les plus faciles à définir et à mesurer, sa signification est complexe. On peut la considérer comme un indicateur d'au moins quatre facteurs très différents, et eux-mêmes composites :

— les phénomènes de maturation et de vieillissement,
— la succession des étapes du cycle de vie,
— les différences entre générations,
— la mortalité différente dans différents groupes sociaux.

Le premier groupe de phénomènes est essentiellement physiologique, avec quelques aspects psychologiques. Sans aller jusqu'à la rigueur de Durkheim, pour qui une explication sociologique ne devrait renvoyer qu'à d'autres phénomènes sociologiques, il faut reconnaître qu'on recourt rarement à cet aspect, dans l'interprétation des corrélations avec l'âge.

C'est que la physiologie elle-même n'est pas indépendante des facteurs sociaux. Non seulement le vieillissement ne se produit pas de la même façon dans des catégories sociales qui vivent différemment, mais dans les rôles sociaux attribués à des personnes d'âges différents, ces aspects physiologiques interviennent : on s'attend à ce qu'une personne âgées soit malade ou fatiguable, par exemple, et ces attentes, liées à la représentation sociale de la vieillesse, retentiront sur la manière dont cette même personne percevra et ressentira son état de santé.

Le cycle de vie, lui, est plus immédiatement explicatif. Dans notre société, la succession des différentes étapes est relativement constante : jeune célibataire, marié sans enfant, marié avec enfants en bas âge, marié avec enfants plus grands ou ayant quitté le foyer, etc. Une telle succession, dont on imagine facilement qu'elle représente une suite de modes de vie très différents, permet de classer sans difficulté la majorité de la population, et l'âge en est un bon indicateur. On peut toutefois se demander si l'augmentation du nombre de divorces, par exemple, ou la diffusion des méthodes contraceptives, ne va pas limiter la généralité de ce cycle traditionnel, rendre sa corrélation avec l'âge plus incertaine, et faire apparaître des étapes moins strictements ordonnées.

Interpréter une corrélation avec l'âge en fonction du cycle de vie est relativement aisé, et il est possible de vérifier empiriquement la validité de cette interprétation. Si, par exemple, on suppose que la diminution du nombre de sorties entre 20 et 35 ans est due à l'accroissement du nombre de personnes mariées, et à la présence d'enfants, on peut vérifier si, à âge égal, les personnes mariées sortent effectivement moins que les célibataires. Ou alors, on peut directement remplacer l'âge par un indicateur plus direct du cycle de vie, en combinant l'état matrimonial, le nombre et l'âge des enfants, et en l'utilisant comme variable explicative. En procédant ainsi, on limite le champ des interprétations possibles, mais il faut prendre garde à ne pas conclure trop vite. Les personnes mariées peuvent sortir moins directement à cause du mariage, par exemple parce qu'elles ont un foyer où elles peuvent rester, parce qu'elles n'ont pas besoin de sortir pour ne pas être seules, etc. ; mais il faut peut-être aussi faire appel à une « sous-culture jeune », ou à des normes sociales spécifiques aux jeunes, qui règlent leur comportement, et qui s'appliquent en premier lieu aux célibataires, les mariés étant censés, quel que soit leur âge, se comporter en « adultes ».

La signification des différentes étapes du cycle de vie, sur ce point entre autres, n'est pas la même dans les différentes catégories sociales. Chez les ouvriers, par exemple, le mariage entraîne une rupture dans le mode de vie, beaucoup plus profonde que dans d'autres catégories. Il en va de même de la façon de percevoir les contraintes liées aux enfants et de s'y soumettre, ou de la manière de vivre après la retraite. Or, de tels phénomènes risquent d'être complètement masqués si on ne considère que la corrélation entre l'âge et un comportement dans l'ensemble de la population. Ce qu'on trouvera ne sera souvent qu'une moyenne effaçant les différences entre les catégories sociales, et donc sans grande signification, si ce n'est purement descriptive.

La définition du cycle de vie en termes de situation familiale est classique, mais elle est loin de recouvrir la totalité de ce qui peut varier avec l'âge. Nous avons vu que les normes et les rôles, aux différents âges, sont différents. La situation professionnelle et économique, elle aussi, évolue au cours de l'existence, ainsi que, par exemple, le type d'intégration sociale.

Mais le problème le plus délicat posé par l'interprétation des variations en fonction de l'âge, c'est celui qui conduit à tenir compte des *effets de génération*, qu'il faut distinguer des évolutions dont nous venons de parler, et qu'on peut toutes rattacher, plus ou moins directement, à l'effet de l'âge proprement dit.

Reprenons, pour préciser, l'exemple des sorties qui diminuent nettement avec l'âge, et ceci dans toutes les catégories sociales, bien qu'à des rythmes différents. L'interprétation de cette décroissance en termes d'âge proprement dit, consiste à supposer que, lorsqu'on avance en âge, des facteurs divers (attrait du foyer, manque d'intérêt pour ce qu'on peut faire hors de chez soi, fatigabilité, etc.) amènent à sortir moins ; ceux qui, actuellement, sont jeunes et sortent beaucoup, subiront la même évolution et, dans quarante ans, se comporteront comme les personnes âgées d'aujourd'hui. Mais on peut aussi supposer que ces mêmes personnes âgées ont été jeunes dans un monde où l'on sortait moins que maintenant, où l'on disposait rarement d'une voiture, où le niveau de vie était plus bas, etc. Ils ne sont donc jamais sortis beaucoup, et ne font actuellement que perpétuer des habitudes acquises dans leur jeunesse. Si cette interprétation

est correcte, dans quarante ans, les personnes âgées sortiront nettement plus que maintenant.

Le problème est central pour l'interprétation de toutes les corrélations avec l'âge : selon qu'il s'agit d'une évolution ou d'un effet de génération, les conséquences aussi bien théoriques que pratiques sont très différentes. Par exemple, si l'on désire prendre des mesures à long terme en faveur des personnes âgées, peut-on se contenter d'étudier leurs besoins actuels, en supposant que ceux qui seront vieux dans vingt ou trente ans seront semblables, pour l'essentiel, à ceux qui le sont aujourd'hui, ayant subi le même vieillissement, ou faut-il admettre qu'ils auront des réactions différentes, propres à leur génération ? Peut-être les différences qu'on observe actuellement entre jeunes et vieux sont-elles dues en premier lieu au fait que ces derniers ont vécu leur jeunesse dans une société très différente de celle où grandirent les jeunes actuels, et qu'ils y ont acquis des habitudes, des représentations, des normes qui guident encore maintenant leurs attitudes et leurs comportements.

Ces deux interprétations, en termes d'âge proprement dit ou en termes de génération, ont évidemment chacune leur part de vérité, mais il faudrait pouvoir préciser dans chaque cas sur quels points c'est l'une ou l'autre qui est la plus vraisemblable. Tant qu'on ne dispose pour cela que d'une enquête unique, effectuée donc à un seul moment bien déterminé, il n'existe aucun moyen rigoureux de trancher : on ne peut qu'avancer des hypothèses plus ou moins argumentées, fondées sur les représentations que nous avons de l'évolution de notre société. Seules des enquêtes identiques, répétées, sur des échantillons comparables, ou sur les mêmes personnes, à intervalles réguliers, et pendant une longue période, la durée d'une génération, par exemple, pourraient permettre de trancher. A notre connaissance, aucune recherche systématique de ce genre n'a été entreprise ; mais l'accumulation d'enquêtes diverses depuis quelques dizaines d'années devrait permettre quelques analyses secondaires, permettant d'avoir une idée de la manière dont les corrélations entre l'âge et divers comportements, ou d'autres caractéristiques, ont varié au cours du temps.

On peut également améliorer la vraisemblance des hypothèses sur ce point en essayant de préciser à quelle composante de l'âge est attribuable la corrélation qu'on cherche à interpréter. Si, par exemple, on peut montrer que cette corrélation est due, en fait, à la variation avec l'âge de la proportion de personnes mariées, et que le comportement des personnes se trouvant dans le même état matrimonial est le même à tous les âges, on pourra exclure un effet de génération et conclure soit qu'on se trouve devant un effet « pur » de l'âge (l'état matrimonial étant considéré comme partie intégrante de l'âge), soit que l'âge, en fait, ne joue aucun rôle (sur le problème de l'interprétation des composantes des variables classiques, cf. 7.3.3.2.).

d) Le sexe

Nous parlerons peu de la troisième de nos variables fondamentales, non que nous en contestions l'importance, mais parce qu'elle pose peu de problèmes de méthode. Même si l'on peut argumenter à partir de cas incontestablement ambigus, ceux-ci restent relativement rares, et en général le classement d'un individu dans un sexe ou l'autre ne pose guère de problèmes. Les travestis, les transsexuels sont peu nombreux, et ne risquent

guère de fausser les enquêtes. Cette situation peut d'ailleur évoluer, et il se pourrait qu'une plus grande tolérance de la société occidentale à leur égard transforme le problème, et exige une nouvelle définition sociale du sexe. Notons que certaines civilisations attribuent un statut particulier aux travestis.

L'interprétation des différences observées entre les sexes selon d'autres variables présente le même type de difficultés que l'interprétation des corrélations avec l'âge : on peut toujours distinguer les composantes physiologiques, psychologiques, sociologiques, économiques. Les rôles masculins et féminins sont très tranchés, même s'ils évoluent, et s'ils varient selon les catégories sociales, et aussi selon les catégories d'âge. Il semblerait que les différences entres les hommes et les femmes soient plus faibles chez les jeunes que chez les personnes plus âgées, et qu'on puisse attribuer ce phénomène à un effet de génération. Mais cela demande encore confirmation.

e) Les relations entre variables classiques

Les principales variables dont nous avons parlé ne sont pas indépendantes les unes des autres, ce qui peut compliquer l'interprétation des corrélations qu'on peut observer entre elles et d'autres variables.

Nous avons vu que la *Catégorie Socio-Professionnelle*, le *revenu* et le *niveau d'instruction* sont fortement corrélés entre eux. Ils entretiennent également des relations complexes avec l'*âge* sur l'ensemble de la population : le revenu commence par croître avec l'âge, puis atteint un palier ou décroît, ces évolutions variant selon la C.S.P.

D'autre part, le *niveau d'instruction* et l'*âge* sont en corrélation inverse. La nature de la relation est très différente de la précédente : le revenu de chaque personne varie effectivement avec l'âge, alors que le niveau d'instruction ne s'abaisse évidemment pas. Il s'agit typiquement d'un effet de génération : la diffusion des enseignements secondaire et supérieur fait qu'un jeune, actuellement, a beaucoup plus de chances d'avoir été au lycée ou à l'université qu'il y a même vingt ans. Sur un échantillon de l'ensemble de la population, la corrélation entre les deux caractéristiques sera forte.

La relation entre le *sexe* et le *revenu* est bien connue, celui des femmes actives est, en moyenne, nettement inférieur à celui des hommes, et ceci par la combinaison de deux mécanismes : d'une part, elles occupent le plus souvent des emplois moins bien rémunérés, et d'autre part, à emploi égal, elles sont souvent moins bien payées. On observe aussi une relation entre le *sexe* et le *niveau d'instruction*, les femmes ayant un niveau d'instruction plus faible, en moyenne, que celui des hommes.

Enfin, notons que l'âge moyen des femmes est supérieur à celui des hommes, ce qui traduit, actuellement, la différence de leurs espérances de vie (il y a cinquante ans cette différence était autre, due surtout aux morts de la première guerre mondiale, en majorité des hommes jeunes). On sait également que les espérances de vie sont différentes selon les C.S.P., ce qui fait que les échantillons de chacune pourront avoir des âges moyens différents.

L'existence de ces multiples relations peut égarer lors de l'interprétation de certaines corrélations. Il n'est pas rare, dans certains résultats d'enquêtes, de voir affirmer qu'un certain comportement diminue avec l'âge, augmente avec le revenu, diminue avec le niveau d'instruction, etc., ces observations étant livrées telles quelles au lecteur, ou in-

terprétées séparément. En fait, elles peuvent être dues à une variable unique, la corrélation avec les autres étant simplement due aux relations générales dont il est question ici. Comme ces relations sont statistiques, et d'importances différentes, à peu près n'importe quelle constellation de corrélations, même apparemment contradictoires, est possible.

Par exemple, on pourra être tenté d'interpréter une corrélation entre l'âge et une consommation par une évolution des goûts, ou par un effet de génération portant sur le style de vie, mais il se peut aussi que ce soit simplement un effet de revenu, ce qu'on vérifiera en cherchant si des personnes ayant le même revenu ont en moyenne la même consommation, quel que soit leur âge. Nous reviendrons plus loin (7.3.3.) sur ces contrôles.

Toutes les corrélations que nous avons signalées sont observables actuellement, dans la population française totale, mais pas nécessairement dans des sous-populations particulières. Par exemple, la différence de niveau d'instruction entre hommes et femmes est moins marquée chez les jeunes et dans certaines catégories sociales ; elle va même jusqu'à être inversée dans certaines régions comme le Nord, où l'absence presque totale d'emplois féminins fait qu'une jeune fille n'a pas intérêt à abandonner ses études trop tôt pour gagner sa vie. Dans l'analyse des résultats de toute enquête, il faut donc commencer par vérifier ce que deviennent ces relations dans l'échantillon particulier étudié, et en tenir compte dans les interprétations.

7.2.3. Les tris à plat et l'estimation des grandeurs

A ce stade, encore très au début, du dépouillement, nous disposons pour chaque sujet de sa valeur sur toutes les variables qui nous intéressent. Certaines ont été directement recopiées sur le questionnaire, d'autres sont le résultat des processus de construction plus ou moins complexes que nous avons vus. Nous pouvons donc connaître la distribution, dans l'échantillon, de toutes ces variables. Dans le jargon de la mécanographie et de l'informatique, c'est ce qu'on appelle les *tris à plat* : pour toutes les valeurs de chaque variable, on indique le nombre de sujets qui présentent cette valeur.

Dans certains cas, relativement peu nombreux, ces distributions à une dimension nous intéressent pour elles-mêmes. Ce sont tous les cas où nous voulons obtenir de l'enquête l'*estimation* de certaines grandeurs, en général des *moyennes* ou des *proportions*. Nous avons vu plus haut (4.1.) pourquoi, en psychologie et en sociologie, il est rare que ces estimations nous intéressent directement. Mais ces cas existent, néanmoins.

Lorsqu'un échantillon est représentatif, une moyenne ou une proportion calculées sur cet échantillon sont des estimations non biaisées des valeurs correspondantes dans la population. Mais, évidemment, le fait que ces estimations ne sont pas biaisées ne veut pas dire que les valeurs trouvées sont exactes. Des erreurs d'échantillonnage sont toujours possibles ; c'est pourquoi il faut toujours associer à une valeur estimée son *intervalle de confiance* : ce sont les limites à l'intérieur desquelles la vraie valeur a 95 % de chances de se trouver (si c'est le seuil qu'on a choisi).

Pour une proportion p, l'intervalle de confiance est donné par :

$$p \pm 1{,}96 \sqrt{\frac{p\,(1-p)}{N}}$$

où N est le nombre d'observations de l'échantillon, et p la proportion observée dans l'échantillon.
Pour une moyenne, l'intervalle de confiance est :

$$m \pm 1{,}96 \frac{\sigma}{\sqrt{N}}$$

où m et σ sont respectivement la moyenne et l'écart-type de l'échantillon.

Ces formules appellent deux remarques. La première, c'est que l'intervalle de confiance, donc la précision de la mesure, ne dépend que de N, la taille de l'échantillon, et non du taux de sondage, lorque la taille de l'échantillon est suffisamment grande. Deuxièmement, le fait que ce soit la racine de la taille qui intervient, et non cette valeur elle-même, rend la précision relativement peu sensible à une variation de l'importance de l'échantillon : il faut quadrupler celui-ci pour que l'intervalle de confiance soit diminué de moitié.

Lorsqu'une enquête ne porte pas sur la population totale, mais sur une sous-population particulière, certains problèmes d'estimations sont, en fait, des problèmes de relation camouflés. Par exemple, si on a pris un échantillon représentatif des lecteurs d'un journal, la composition socio-professionnelle de cette sous-population peut être étudiée directement. D'un point de vue descriptif, c'est suffisant. Mais on risque, en s'en tenant là, de commettre des erreurs graves. En effet, on ne peut vraiment tirer des conclusions de la connaissance de cette composition que par comparaison avec les caractéristiques correspondantes des non-lecteurs, ou, ce qui revient au même, de la population totale.

Supposons que l'on veuille décrire cette sous-population au moyen d'une variable à trois modalités, donnant la distribution suivante :

A	60 %
B	30 %
C	10 %
Total	100 %

Ce qu'on constate immédiatement, c'est que, parmi les lecteurs de ce journal, la catégorie A est nettement majoritaire, et que C est très peu représentée. Mais on aurait tort d'en inférer que ce journal a un succès particulier auprès des A, et qu'il n'intéresse pas les C. Une telle conclusion ne serait possible que par référence à une population totale où, par exemple, les trois catégories seraient également représentées. Mais supposons que la composition de celle-ci soit :

A	70 %
B	25 %
C	5 %

La comparaison des deux distributions nous amène à une conclusion inverse de la précédente : les A sont moins représentés parmi les lecteurs que dans la population totale, alors que les C le sont plus. En fait, c'est dans la catégorie C que le journal peut être considéré comme ayant le plus de succès, mais comme elle est peu nombreuse dans la population, elle ne peut fournir que peu de lecteurs. Nous reviendrons en détail sur ces problèmes à propos de l'analyse des tableaux croisés (voir 7.3.5.). Pour le moment, ce qu'on peut retenir de cet exemple (fictif), c'est qu'il faut prendre garde aux risques de comparaisons implicites, et que ce qui peut apparaître comme un simple problème d'estimation cache, en fait, l'analyse de relations. En effet, ce qui est en jeu ici, ce qui interviendra dans le discours qu'on tiendra sur ces résultats, ce n'est pas seulement l'estimation, c'est aussi la comparaison des lecteurs et des non-lecteurs.

Supposons que les lecteurs constituent 20 % de la population totale. Nous pouvons, à partir des données précédentes, reconstituer le tableau complet permettant la comparaison :

	Lecteurs	Non-lecteurs	Total
A	60 %	72 %	70 %
B	30 %	24 %	25 %
C	10 %	4 %	5 %
	100 %	100 %	100 %

On voit apparaître tout à fait clairement la faiblesse de proportion de lecteurs dans la catégorie A.

7.3. LES RELATIONS

7.3.1. Définitions

Les hypothèses qu'on peut vérifier au moyen d'une enquête, comme toute hypothèse scientifique, s'expriment nécessairement sous la forme de *relations* entre deux ou plusieurs variables, dans la population totale ou dans des sous-populations particulières. La mise en évidence et l'interprétation de relations constituent donc l'essentiel de l'analyse des résultats d'une enquête.

Nous emploierons ici le terme *relation* dans son sens le plus général, couvrant à la fois :

— les *corrélations* entre variables quantitatives ou ordinales,
— les *associations* entre variables qualitatives (nominales),
— les *différences* entre groupes.

Ce dernier cas n'est en général pas présenté comme une relation, mais formellement c'est bien de la même chose qu'il s'agit. Supposons que nous observions une différence entre les distributions d'une certaine variable X chez les hommes et chez les femmes. On pourra exprimer cette constatation en disant soit que les hommes et les femmes

diffèrent, en moyenne, du point de vue de X, soit qu'il y a une *relation* entre X et le sexe. Dans le premier cas, on considère que le sexe distingue, définit, deux groupes, dans le second que c'est une variable à deux modalités ; mais les observations décrites sont les mêmes, ainsi que ce qu'on peut en conclure. Les tests statistiques permettant de vérifier ces deux formes de l'hypothèse (test de comparaison de moyennes, ou test de Student dans le premier cas, coefficient de corrélation bisérial dans l'autre) sont rigoureusement équivalents.

Presque toutes les techniques de vérification de l'existence d'une relation, ou de mesure de la force de celle-ci, supposent des relations formellement symétriques, où donc les deux variables jouent exactement le même rôle. C'est le cas de différents coefficients de corrélation (de Bravais-Pearson, de Spearman, de Kendall) ou d'association. Pourtant comme nous le verrons plus loin, des relations asymétriques sont formellement possibles (implication), et d'autre part les hypothèses sont souvent exprimées sous la forme de relations asymétriques (causalité, conditions nécessaires ou suffisantes, etc.). Il y a donc un hiatus entre la nature des hypothèses et les instruments disponibles pour les vérifier.

Une autre limite des mesures de corrélation usuelles, c'est qu'elles portent sur des relations *monotones*, le plus souvent *linéaires*. Certes, il est rare que, dans les sciences sociales, on soit en mesure de spécifier suffisamment une hypothèse pour faire des prédictions sur la forme que devrait prendre la relation. On se contente de dire qu'une grandeur doit croître ou décroître en fonction d'une autre, sans autre précision. Lorsqu'on utilise un coefficient comme celui de Bravais-Pearson, qui mesure une corrélation linéaire, cela peut amener à sous-estimer la force de la relation lorsque celle-ci, en fait, a une forme non linéaire, même si, conformément à l'hypothèse, elle reste monotone.

Pour le moment, nous n'avons parlé que des relations entre *deux* variables. Or, bien des hypothèses amènent à prendre en considération des relations plus complexes, faisant appel à trois variables ou plus. L'analyse et l'interprétation deviennent très vite plus compliquées, dans la mesure où on ne dispose pas d'un langage commode pour en parler. C'est pourquoi, très souvent, on décomposera une relation entre trois variables en un plus grand nombre de relations entre deux variables, qu'on interprétera simultanément (voir 7.3.3.).

7.3.2. L'analyse des relations binaires

2.3.2.1. Le langage des relations

Le langage utilisé habituellement pour parler des relations qu'on observe est très varié.

Au lieu de dire qu'il existe une corrélation entre X et Y, on dira par exemple :

— X est *fonction* de Y (ou Y est fonction de X),
— X *varie avec* Y, ou *comme* Y,
— X et Y sont *liés*,
— X *évolue avec* Y.

En introduisant des éléments d'interprétation plus poussés, on pourra aussi dire :

— X *dépend* de Y,
— X *cause* Y,
— X *entraîne* Y,
— X *implique* Y,
— X *détermine* Y,
— X *est une condition de* Y.

Les connotations de ces diverses expressions sont très différentes mais elles recouvrent toutes la même observation : les valeurs de Y ne sont pas les mêmes selon les différentes valeurs de X auxquelles elles sont associées. Dans ces conditions, qu'est-ce qui justifie d'employer une expression de préférence à une autre ?

Remarquons d'abord que, dans la plupart de ces formulations, X et Y ne jouent pas le même rôle. En effet, si certaines sont nettement symétriques (X et Y sont liés), la plupart peuvent être considérées comme asymétriques (X cause Y, X entraîne Y, X est une condition de Y), d'autres étant ambiguës (X dépend de Y, X est fonction de Y). Dans quels cas dira-t-on que « X détermine Y » plutôt que « Y détermine X » ?

Enfin, certaines expressions sont simplement descriptives (X et Y sont liés), alors que d'autres constituent déjà des interprétations (X cause Y).

On pourrait considérer que toutes ces expressions ne sont qu'autant d'abus de langage, et que seule la proposition « X et Y sont corrélés » est légitime, a une signification précise. Mais on s'enferme alors dans une description purement empirique, interdisant toute explication de ce qui est observé. Si l'on veut dépasser la simple constatation, il faut préciser plus clairement les différentes significations possibles du mot « relation », ainsi que les diverses interprétations qu'on peut en donner. Pour cela, il est nécessaire, une fois de plus, de distinguer le langage des observables de celui de la théorie même si, souvent, la théorie en question se réduit à quelques présupposés mal explicités.

Dans ce qui suit, nous commencerons par examiner comment décrire des relations entre observables, et nous verrons que, formellement, on n'en distingue qu'un très petit nombre, qui serviront ensuite de termes de référence pour l'analyse des relations effectivement observées. Puis, nous essaierons de clarifier ce qu'on entend par « interprétation » d'une relation. Lorsqu'on dispose d'une théorie bien constituée, interpréter une relation observée revient en général à montrer qu'elle est un cas particulier d'application d'une loi générale. Mais quand, ce qui est le cas le plus fréquent dans les enquêtes, cette théorie est insuffisante ou même manque complètement, nous en sommes réduits à faire appel aux connaissances tirées de notre vie quotidienne, éventuellement éclairées et précisées par ce que nous pouvons connaître de psychologie ou de sociologie.

7.3.2.2. Les relations entre observables

Commençons par nous situer au niveau des observables, et des relations qu'on peut mettre en évidence entre eux. Considérons d'abord un cas particulièrement simple, celui des diverses relations possibles entre deux variables dichotomiques, présence ou ab-

sence d'une certaine caractéristique, que nous noterons, comme cela se fait souvent, X et $\overline{X}$, Y et $\overline{Y}$.

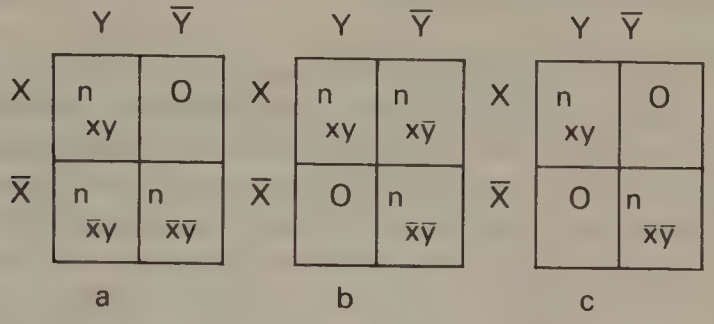
Pour analyser les relations qui peuvent exister entre elles, on se réfèrera à 3 « cas purs », qu'on peut définir formellement :

a. X implique Y (X est une condition suffisante de Y),

b. Y implique X (X est une condition nécessaire de Y),

c. X et Y s'impliquent réciproquement (X est condition nécessaire et suffisante de Y).

Ces trois relations se traduisent par les tableaux de fréquence correspondants :



A ces trois relations fondamentales, qui nous serviront désormais de référence, on peut en ajouter trois autres, obtenues en remplaçant Y par $\overline{Y}$, ce qui donne des relations d'exclusion ou d'incompatibilité.

Habituellement, au lieu des zéros, on observera des fréquences faibles. On est donc obligé de recourir à des critères, par exemple des tests statistiques, qui nous indiqueront des seuils au-delà desquels la fréquence dans les cases théoriquement vides devient trop élevée pour accepter qu'on se trouve en présence d'une approximation de la relation pure correspondante [1]. Mais les tests statistiques courants (test de χ^2 par exemple) ne permettent pas de distinguer les trois cas ci-dessus, puisqu'ils vérifient simplement l'hypothèse d'indépendance des deux variables, sans que l'hypothèse alternative soit précisée. Or, les trois tableaux ci-dessus s'écartent de cette hypothèse nulle : une valeur significative de χ^2 permettra de rejeter l'hypothèse nulle, donc de conclure à l'existence d'une relation, mais sans préciser laquelle. Simplement, à distributions marginales identiques, la valeur de χ^2 sera plus élevée dans le cas de l'implication réciproque que dans les deux autres. Toutefois, devant un cas particulier, cette constatation est insuffisante pour conclure.

Telles qu'elles sont formulées de façon verbale, les hypothèses psychologiques ou sociologiques distinguent d'ailleurs mal ces trois relations. Par exemple, lorsqu'on dit que l'appartenance à une catégorie sociale (X) favorise, ou entraîne (termes du langage de la théorie, non des observables) une certaine attitude, cela peut vouloir dire que *tous* les membres de X présentent l'attitude Y (relation a), que *seuls* les membres de X présentent Y (relation b), ou les deux à la fois (relation c). En principe, les conclusions et les interprétations qu'on pourra tirer de chaque cas devraient être différentes. Mais, en pratique, comme X n'est très vraisemblablement pas l'unique déterminant de Y, on

1. Le coefficient de reproductibilité, ou les autres coefficients utilisés en analyse hiérarchique, constituent de tels critères spécifiques pour les relations d'implication a ou b.

se contente de constater que la proportion des Y parmi les X est supérieure à leur proportion parmi les $\overline{X}$, formulation générale compatible avec les trois relations, et avec des relations intermédiaires, affaiblies, qui s'exprimeraient, si l'on voulait les formuler de façon explicite, en termes probabilistes, et non simplement en approximations du tout ou rien. Nous reviendrons, à propos de l'interprétation des relations, sur les significations qu'on peut attribuer aux écarts par rapport aux relations « pures ».

7.3.2.3. L'INTERPRÉTATION DES RELATIONS

Si nous passons maintenant au langage de la théorie ou, pour être moins ambitieux, au langage des interprétations, la variété des termes employés devient beaucoup plus grande. Lorsque nous observons une relation, quelle qu'elle soit, entre deux variables X et Y, les *interprétations* suivantes sont possibles, sans que cette liste soit exhaustive :

a. X et Y sont *tous deux en corrélation avec une troisième variable*, et leur relation apparente est due uniquement à celle-ci.

b. X et Y sont deux *indicateurs d'une même variable intermédiaire*. Par exemple, on pourra interpréter la corrélation entre deux opinions en disant qu'elles sont deux expressions de la même attitude sous-jacente.

c. X et Y sont deux *effets d'une même cause*. Par exemple, la fréquentation des musées et les voyages à l'étranger peuvent tous deux être attribués à un revenu élevé.

Pour ces trois cas, la relation entre X et Y n'est qu'apparente, et due à une troisième variable. Dans les cas suivants, on peut la considérer comme « vraie », et donc susceptible d'être interprétée pour elle-même :

d. X et Y sont *fonctionnellement complémentaires*, comme la possession d'un électrophone et celle de disques.

e. X est un *stimulus* ou, plus largement, un *événement*, et Y la *réaction* à cet événement.

f. X est une *caractéristique individuelle* (par exemple l'âge, le sexe, l'appartenance à une catégorie sociale, à un groupe religieux, etc.) et Y un *comportement* ou une *attitude*, ou un autre trait de même nature.

g. X est une *disposition* (attitude, aptitude, trait de personnalité, etc.) et Y un *comportement*.

h. X est une *condition* (nécessaire ou suffisante) d'un certain *effet* Y. Par exemple, certaines connaissances peuvent être nécessaires pour réaliser une tâche donnée.

Dans cette liste d'interprétations possibles, certaines (a-d) sont symétriques, en ce sens que les deux variables X et Y y jouent le même rôle ; les autres (e-h) sont asymétriques, les deux variables y jouent des rôles différents : on parlera souvent de *variable explicative* et de *variable expliquée*, distinction qu'on peut rapprocher de celle que la psychologie expérimentale fait entre variables *indépendante* et *dépendante*. Dans les enquêtes, il n'est pas possible de manipuler les variables, et donc la notion d'indépendance, dans ce sens particulier, n'est pas transposable. Mais nous verrons que certaines variables peuvent être considérées comme antérieures à d'autres, logiquement ou temporellement.

Parmi les interprétations asymétriques, la plupart (e-g) peuvent être considérées comme des formes diverses de *causalité*. Ce n'est pas ici le lieu d'entreprendre une discussion épistémologique de ce concept. Rappelons simplement que ce n'est en aucun cas une relation observable, mais une interprétation, et que les limites de l'emploi du terme sont très vagues. Actuellement, parmi les chercheurs en sciences sociales, on en trouve qui refusent d'utiliser le terme (ou qui estiment que la transposition d'une conception de la causalité issue des sciences physiques n'a pas de sens pour la compréhension des phénomènes humains), alors que d'autres l'appliquent à toutes les interprétations asymétriques [1].

Il ne faut pas confondre l'asymétrie dans l'*interprétation* des relations avec l'asymétrie formelle dont il a été question précédemment. Il n'y a aucun lien entre le caractère symétrique ou non de la relation observée et celui de l'interprétation, qui se fonde sur la représentation qu'on se fait du mécanisme qui relie les deux variables.

Pour préciser ce point, prenons la proposition très générale « X cause Y » qui, comme nous l'avons dit, appartient sans ambiguïté au langage de la théorie, et qui est, par définition, asymétrique. Trois spécifications, correspondant aux trois relations formelles ci-dessus, sont possibles :

	Relation observée	Interprétation
a.	X implique Y	X cause toujours Y, mais Y peut avoir d'autres causes que X.
b.	Y implique X	Seul X peut causer Y, mais il ne le cause pas toujours ; il doit être associé à d'autres variables pour produire Y.
c.	Implication réciproque entre X et Y	X cause toujours Y, et X est la seule cause possible de Y.

On voit que la « loi scientifique », celle qui « explique » le mieux Y, c'est la relation symétrique c, qui établit une correspondance complète entre cause et conséquence. Les deux autres relations ne sont qu'approximatives, puisqu'elles ne disent rien des autres conditions d'apparition de Y. La relation *a* indique que d'autres causes de Y sont possibles (X *ou* Z_1, *ou* Z_2...causent Y) ; la relation *b* que d'autres sont nécessaires, en plus de X (X *et* Z_1, *et* Z_2... causent Y). Les cas où l'on observe X et $\overline{Y}$, $\overline{X}$ et Y, montrent l'existence d'autres variables Z, non identifiées, et probablement réparties de façon aléatoire.

Lorsque nous obtenons empiriquement une approximation de la relation symétrique *c*, il faudrait probablement l'interpréter comme un mélange des situations *a* et *b*, c'est-à-dire comme l'indication qu'il y a d'autres conditions à l'apparition de Y, en plus de X. Cela amène à rechercher quelles sont les autres variables, et à construire des *schémas de causalité* qui tiennent compte de l'influence simultanée de plusieurs facteurs. C'est là l'idée de base de l'analyse de causalité, sur laquelle nous reviendrons.

1. C'est le cas des chercheurs qui ont développé les techniques dites d'analyse de la causalité, et qui essaient de donner un statut plus clair à cette notion.

S'il n'y a pas correspondance entre le caractère symétrique ou asymétrique d'une relation et celui de son interprétation, comment peut-on choisir celle-ci ? Comment décidera-t-on quelle est la variable antécédente et quelle est la conséquente, la cause et l'effet, la variable explicative et la variable expliquée ? Pourquoi dira-t-on, même au niveau de l'analyse formelle, que X est une condition nécessaire de Y, et pas que Y est une condition suffisante de X ?

Dans l'expérimentation en laboratoire, la possibilité de manipuler les conditions permet de distinguer variables indépendantes et variables dépendantes. Cette distinction n'a pas, au sens strict, son équivalent dans les enquêtes : formellement toutes nos variables sont des réponses recueillies dans les mêmes conditions. Il nous faut alors nous appuyer sur le processus qu'on peut imaginer derrière la relation, processus dont elle serait l'aboutissement ou l'expression ; il nous faut donc une théorie, même sommaire.

En l'absence d'une telle théorie explicite, on est obligé de s'appuyer sur une théorie intuitive de l'antériorité, logique ou chronologique : l'appartenance à une catégorie sociale est antérieure à une attitude, un événement est antérieur à ses conséquences. Mais l'antériorité n'est pas toujours facile à établir. Certaines attitudes, par exemple, peuvent faciliter la mobilité sociale : dans ce cas ce sont elles qui seront antérieures à l'appartenance à une certaine catégorie. De même, la théorie de la dissonance cognitive juge trop simple l'idée que c'est l'attitude qui détermine certains comportements : l'inverse est également possible, certains comportements peuvent entraîner une modification des attitudes.

Ainsi, lorsqu'une relation est interprétée de façon asymétrique, c'est-à-dire lorsqu'on attribue aux deux variables des rôles différents dans le processus explicatif, le fait qu'elle soit formellement asymétrique indique l'intervention simultanée d'autres facteurs que ceux pris en considération, et donc d'une causalité complexe. Lorsque l'interprétation est symétrique, la signification de la relation d'implication est très différente. Nous avons déjà rencontré ce cas : l'implication est la relation de base de l'analyse hiérarchique. Nous avons vu qu'il est possible, dans ce cas, c'est-à-dire lorsque les variables sont de même nature, d'interpréter la relation soit comme l'expression d'une nécessité logique, soit comme une cohérence de type social (cf. 7.2.1.2.).

Jusqu'à présent, nous avons raisonné sur le cas particulier de variables dichotomiques, dont les deux modalités représentaient la présence ou l'absence d'une certaine caractéristique. Bien que, formellement, par des regroupements adéquats, on puisse très souvent ramener des cas très divers à celui-ci, il est nécessaire de voir ce que devient notre analyse dans d'autres situations, en particulier dans le cas le plus différent, celui de deux variables quantitatives. Tout ce que nous avons dit sur les interprétations possibles reste valable dans ce cas. En revanche, il est plus difficile de mettre en évidence des relations asymétriques entre les observations. Certes, conceptuellement, c'est tout à fait possible : on peut admettre qu'une valeur élevée de X est une condition nécessaire, mais non suffisante, d'une valeur élevée de Y. Par exemple, ce pourrait être le cas de la relation entre le revenu et une certaine consommation : les personnes de revenu bas consommeraient peu, alors que celles ayant un revenu élevé pourraient consommer

beaucoup, mais ne le feraient pas nécessairement toutes. Cela se traduirait non seulement par une augmentation de la consommation moyenne en fonction du revenu, mais aussi par une variance plus forte de la consommation dans les catégories à revenu élevé.

7.3.3. Les relations entre trois variables

7.3.3.1. La description des relations

Dans le chapitre précédent, nous n'avons pris en considération, au départ, que des relations entre deux variables. C'est d'ailleurs ainsi que se présentent en général les premiers dépouillements d'une enquête, après les simples tris à plat. L'examen de ces relations peut déjà permettre de vérifier certaines hypothèses et peut aussi en suggérer d'autres ; on peut également se poser des problèmes quant à la signification de ces relations observées, se demandant si elles représentent des liaisons réelles, ou ne sont que des artefacts.

Nous avons vu que l'interprétation de ces relations binaires devait souvent faire appel à d'autres variables, les phénomènes étudiés ne pouvant s'expliquer par un seul facteur. D'autre part, à propos des relations entre les variables explicatives classiques, nous avons vu que, puisqu'elles sont toutes intercorrélées, une relation entre un phénomène quelconque et l'une d'entre elles peut n'être qu'illusoire, et simplement la conséquence que ces variables sont toutes deux liées à une troisième. Cette situation est très fréquente dans les situations réelles qu'étudient les enquêtes, où il est difficile de rendre toutes les variables pertinentes indépendantes, comme on essaie de le faire en laboratoire. Enfin, on peut supposer que la relation « réelle » entre deux variables peut ne pas être la même dans différentes sous-populations définies par les modalités différentes d'une troisième. Dans ce cas la relation binaire qu'on observera dans l'échantillon total pourra être la résultante du mélange de sous-populations diverses, et n'avoir aucune signification en elle-même.

Pour toutes ces raisons, on est amené à étudier les relations entre trois variables, ou même un nombre élevé. Pour le moment, nous nous bornerons au cas de trois variables, qu'il est possible de traiter par la seule analyse des tableaux croisés. Lorsque le nombre de variables augmente, cette saisie directe devient vite impossible, et il faut recourir à des méthodes plus formalisées, comme l'analyse de causalité.

En termes très généraux, on peut dire que l'introduction d'une troisième variable, ou d'un plus grand nombre, peut viser deux objectifs principaux, qui d'ailleurs ne s'excluent pas :

— Mettre en évidence des causalités complexes, faisant intervenir plusieurs variables.

— Eliminer l'influence d'autres variables, pour mettre en évidence la relation « vraie » entre les deux qui nous intéressent.

Le deuxième objectif vise à recréer, à posteriori, des sous-échantillons du même type que ceux qu'on aurait au laboratoire, où les facteurs susceptibles d'influencer les phénomènes étudiés sont rendus indépendants, ou distribués de façon aléatoire entre les groupes expérimentaux.

Analyser directement, et dans sa globalité, un tableau à triple entrée, même si chaque variable n'a que deux valeurs, n'est pas facile. En fait, on procède toujours en constituant différents tableaux à deux dimensions, et en les comparant, ou en traçant les courbes correspondantes.

Lorsqu'on veut étudier la relation entre trois variables, X, Y et Z, on examine donc comment se présente la relation entre X et Y dans les sous-échantillons définis par chacune des modalités de Z. Ce faisant, on a déjà effectué un choix, puisque Z joue un rôle différent de celui de X et de Y. Nous verrons plus loin à quoi correspond cette différence, et les raisons qu'on peut avoir de distinguer une des variables plutôt qu'une autre.

Commençons par examiner formellement les différents cas possibles. Supposons d'abord qu'entre chacune des trois variables on ait observé des relations deux à deux. Nous constituons maintenant des sous-échantillons correspondant aux deux modalités Z et $\overline{Z}$, croisons X et Y dans chaque sous-échantillon. Plusieurs cas sont alors possibles :

a) La relation entre X et Y est la même dans chaque sous-échantillon, et dans l'échantillon total. On peut en conclure que la relation observée n'est pas due à l'influence de Z. On a donc maintenant plus de raisons de supposer que cette relation est « vraie », c'est-à-dire qu'elle est l'expression d'un processus qui relie directement X à Y, et qu'on peut interpréter selon une des manières que nous avons vues ci-dessus. En principe, pour qu'une telle interprétation soit légitime, on ne peut pas se contenter de constater que la relation n'est pas attribuable à Z : il faudrait aussi éliminer toutes les autres hypothèses relatives aux autres variables susceptibles de rendre compte de la relation. Bien qu'on ne puisse jamais éliminer complètement toutes les hypothèses possibles, le fait d'avoir ainsi vérifié que les variables Z, Z', Z'' ne peuvent pas rendre compte de la relation entre X et Y permet d'attribuer à celle-ci une valeur plus grande.

b) La relation entre X et Y n'apparaît dans aucun des sous-échantillons. Cela indique que celle qui avait été observée dans l'échantillon total était due uniquement à l'influence de Z, qui est donc le véritable facteur explicatif. Si l'on croise Z et X dans des sous-échantillons définis pour chaque modalité de Y (ou Z et Y pour chaque X), on retrouvera les mêmes relations que dans l'échantillon total.

On peut prendre comme exemple permettant de mieux saisir la différence entre les cas a et b deux consommations (X et Y), et le revenu (Z). Si la relation entre X et Y se maintient dans chaque catégorie de revenu (cas a), on peut interpréter ce résultat en disant qu'elles sont deux expressions d'un même goût, ou deux éléments d'un même mode de vie, ce goût ou ce mode de vie pouvant être liés au revenu. Si, au contraire, la relation disparaît dans chaque sous-échantillon (cas b), on en conclura qu'il s'agit simplement de deux consommations qui augmentent chacune avec le revenu, mais sans lien spécifique entre elles. Dans le premier cas, ce sont les mêmes personnes qui consomment à la fois X et Y, la proportion de ces personnes croissant avec le revenu ; dans le second, ce pourront être des personnes différentes, bien que toutes caractérisées par un revenu élevé.

Prenons un exemple numérique fictif. Dans l'ensemble de l'échantillon, supposons qu'on ait trouvé les trois relations suivantes :

	X	$\overline{X}$	
Y	100	60	160
$\overline{Y}$	85	135	220
	185	195	380

	X	$\overline{X}$	
Z	50	150	200
$\overline{Z}$	135	45	180
	185	195	380

	Y	$\overline{Y}$	
Z	40	160	200
$\overline{Z}$	140	60	180
	160	220	380

Les trois variables sont corrélées. Si c'est la relation entre X et Y qui nous intéresse, on peut se demander si elle n'est pas due à Z, puisque Z est lié à la fois à X et à Y. On constitue donc les deux tableaux suivants, qui sont la décomposition du premier tableau ci-dessus, en distinguant Z et $\overline{Z}$:

	X	$\overline{X}$	
Y	100	60	160
$\overline{Y}$	85	135	220
	185	195	380

Z

	X	$\overline{X}$	
Y	40	0	40
$\overline{Y}$	10	150	160
	50	150	200

$\overline{Z}$

	X	$\overline{X}$	
Y	120	0	120
$\overline{Y}$	15	45	60
	135	45	180

Il faut noter que les deux tableaux ne se déduisent pas directement du précédent ; c'est empiriquement, par des tris supplémentaires (d'abord en séparant Z et $\overline{Z}$, puis en croisant X et Y dans chaque sous-échantillon ainsi constitué), qu'ils sont construits.

Dans cet exemple, on voit que, dans chacun des deux sous-échantillons, on observe la même association positive entre X et Y. On peut donc en conclure à l'existence d'une relation entre X et Y, ou, plus prudemment, que la relation entre elles n'est pas due à Z. On acceptera la réalité d'une telle relation s'il n'y a pas d'autres variables parasites susceptibles d'expliquer la corrélation.

Supposons maintenant que, avec les mêmes relations dans l'ensemble de l'échantillon, la décomposition selon Z donne les résultats suivants :

Z

	X	$\bar{X}$	
Y	10	30	40
$\bar{Y}$	40	120	180
	50	150	200

$\bar{Z}$

	X	$\bar{X}$	
Y	90	30	120
$\bar{Y}$	45	15	60
	135	45	180

Il y a maintenant indépendance entre X et Y. La relation qu'on pouvait observer dans l'échantillon total était due uniquement à la présence de Z. Selon le cas, on interprétera ce résultat soit en disant qu'il n'y a pas de relation réelle entre X et Y, que la corrélation entre elles n'était qu'apparente, soit en disant que la relation entre X et Y est bien réelle, mais qu'elle s'explique par la présence de Z, qui constitue une maille d'une chaîne causale entre X et Y. Nous reviendrons sur ces interprétations plus loin (cf. 7.3.3.2.).

La comparaison directe de tableaux croisés n'est pas toujours aisée. Elle peut être facilitée par une représentation graphique, surtout, mais pas seulement, lorsque X et Y ont plus que deux valeurs. On reporte sur le même graphique XY les valeurs moyennes de Y pour chaque X (ou l'inverse), et ceci pour chaque valeur de Z. On obtient ainsi, sur le même graphique, une courbe pour chaque sous-échantillon, et la comparaison entre elles est immédiate (fig. 11).

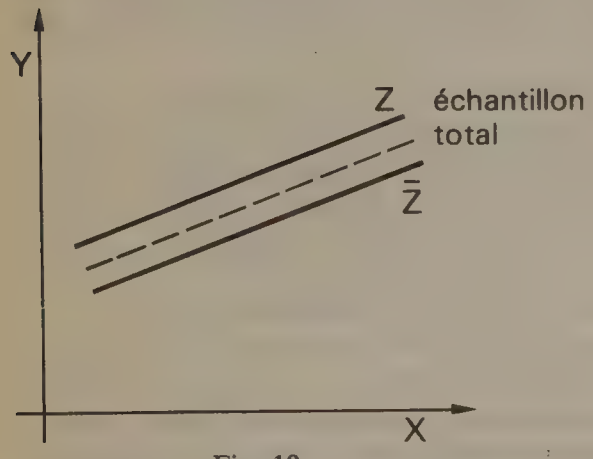


Fig. 10

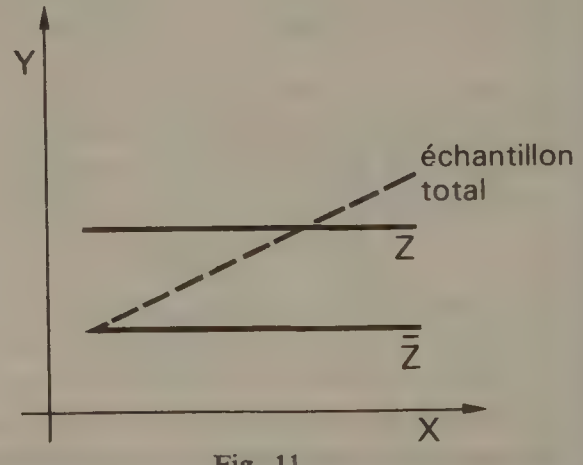


Fig. 11

Le premier graphique correspond au cas a, celui où la relation se maintient dans Z et dans $\bar{Z}$, le second cas b, celui où elle disparaît.

c) La relation entre X et Y n'est pas la même dans chacun des sous-échantillons. Dans ce cas, quelle que soit la relation apparente dans l'échantillon total, elle n'a pas d'autre intérêt que simplement descriptif, puisqu'elle est la résultante de la sommation

d'observations hétérogènes. En particulier, si les relations dans deux sous-échantillons sont inverses, il est possible que toute relation disparaisse au niveau de l'échantillon global. Ce cas est particulièrement important à relever, car il indique que, lorsqu'on soupçonne l'influence d'une troisième variable, l'absence de relation peut être aussi illusoire que sa présence.

Dans des cas entrant dans cette catégorie, on ne peut plus parler, dans l'absolu, de relation entre deux variables, mais seulement de *relations conditionnelles* à une troisième : la relation entre X et Y dépend de la valeur de Z. C'est quelque chose d'analogue à l'*interaction* dans l'analyse de variance ; cette notion est indispensable pour la description de relations entre trois variables, mais elle demande à être spécifiée dans chaque cas. En effet, l'existence d'une interaction peut se traduire de manières très diverses : Y et X sont indépendants dans Z et en corrélation dans $\bar{Z}$ (fig. 12, C_1), la relation peut être inverse dans les deux cas (C_2), les deux courbes peuvent avoir des formes, ou simplement, s'il s'agit de droites, des pentes différentes (C_3, C_4). Lorsque Z a plus de deux modalités, les interactions peuvent évidemment devenir plus complexes.

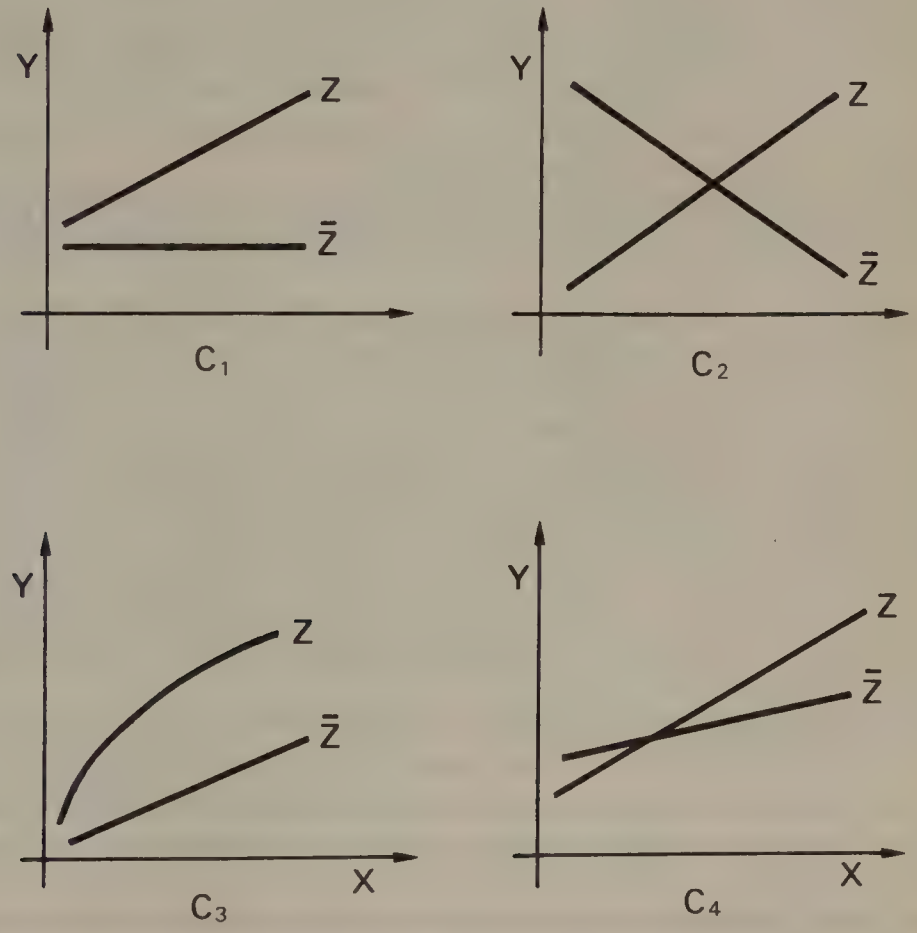


Fig. 12

7.3.3.2. L'INTERPRÉTATION DES RELATIONS ENTRE TROIS VARIABLES

Elle dépend des rôles que l'on fait jouer à chacune, et des hypothèses qu'on a fait à leur propos.

Considérons d'abord le cas où *la relation entre X et Y disparaît*, ou du moins s'atténue fortement dans chacune des modalités de Z. On peut dire alors que Z *explique* la relation entre X et Y. Mais cette notion d'explication reste vague, et renvoie elle-même à plusieurs interprétations possibles.

a) On interprète cette disparition de la corrélation en disant que *la corrélation entre X et Y n'était qu'apparente*, due uniquement à la présence de Z.

On admettra alors que X et Y sont, en réalité, indépendants, et donc qu'il n'y a pas lieu de chercher un mécanisme qui rende compte directement de leur liaison, autre que l'influence de Z. Nous avons vu plus haut un exemple d'une telle situation, lorsque nous avons cité l'exemple de deux consommations dont la liaison était « expliquée » par le revenu. Il s'agissait dans ce cas du lien entre deux variables de même nature, et le problème qui se posait était de savoir, devant la constatation d'une relation entre elles, s'il fallait l'attribuer à une variable intermédiaire (le goût, le mode de vie, etc.) ou si une variable observable pouvait en rendre compte. Si la corrélation disparaît lorsqu'on « contrôle » le revenu (c'est-à-dire lorsqu'on constitue des catégories homogènes du point de vue du revenu), cela indique qu'il n'y a pas lieu de faire appel à d'autres variables explicatives. En revanche, on peut encore se poser le problème d'« expliquer » les relations des deux consommations avec le revenu.

La même démarche s'applique aussi lorsqu'un phénomène (une variable) à expliquer se trouve en corrélation avec deux variables explicatives elles-mêmes corrélées entre elles.

Il est très fréquent de lire, dans des compte rendus d'enquêtes, des affirmations du genre : « Les personnes qui présentent le comportement X sont plus jeunes, ont des revenus plus élevés, sont plus souvent des hommes, habitent plus souvent dans les grandes villes, ont plus souvent des enfants, que ceux qui ne le présentent pas. » Cette énumération de facteurs explicatifs est très ambiguë. Elle peut suggérer à première lecture que le comportement X se trouvera surtout dans une catégorie caractérisée *à la fois* par le fait d'être jeune, riche, etc. Mais il se peut également que chaque facteur intervienne seul, ou encore que ces résultats ne soient que des conséquences de l'existence de corrélations entre toutes ces variables. Il est donc indispensable de ne pas se contenter de cette suite de corrélations entre X et chacune des autres variables, mais de considérer celles-ci simultanément.

Prenon le cas de deux variables explicatives seulement, par exemple l'âge et le niveau d'instruction Nous savons que ces deux facteurs sont liés dans la population totale : une des deux corrélations avec le comportement en question peut donc n'être qu'un artefact. On examinera alors ce que devient, par exemple, la relation entre le comportement et l'âge dans chaque catégorie de niveau d'instruction, ou l'inverse. Si dans un des cas la corrélation disparaît, on pourra en conclure que c'est l'autre facteur qui est un déterminant du comportement étudié.

Par exemple, de nombreuses pratiques culturelles varient avec l'âge, mais on constate que, dans des groupes homogènes du point de vue du niveau d'instruction, cette relation disparaît. On en conclut que la « vraie » relation, celle sur laquelle on peut réfléchir et qu'on peut interpréter, utiliser pour des prévisions, etc., est la relation avec le niveau d'instruction.

b) On interprète Z comme étant une des *composantes* de X, celle qui est responsable de la relation entre X et Y.

Admettons que X soit l'âge, Y un comportement qui décroît avec l'âge, la fréquence des sorties par exemple. On peut considérer que cette relation n'est pas suffisamment explicative en elle-même. En effet l'âge, nous l'avons vu, est une variable complexe, qui regroupe un grand nombre de caractéristiques très différentes, caractéristiques que nous appellerons des *composantes*. Interpréter une corrélation avec l'âge c'est bien souvent préciser laquelle de ces composantes intervient. En général, cette interprétation reste hypothétique : on supposera, par exemple, que si l'on sort moins lorsqu'on avance en âge, c'est parce que la probabilité d'être marié augmente, et que ce sont les célibataires qui sortent le plus. Cette interprétation est vérifiable empiriquement, il suffit de voir si, dans chaque catégorie d'âge, il y a une différence entre les célibataires et les mariés, ou, inversement, si la corrélation avec l'âge disparaît dans le sous-groupe des célibataires et dans celui des mariés. Si c'est le cas, on pourra dire qu'on a expliqué la relation avec l'âge, en précisant à laquelle des composantes de l'âge est due cette relation.

La différence avec le cas précédent tient à la nature de la relation entre l'âge et la situation matrimoniale qui, dans notre société, peut être considérée comme moins contingente que la liaison entre l'âge et le niveau d'instruction. On acceptera que le changement de l'état matrimonial fasse partie de l'âge, en soit une composante. Certes, on peut encore se demander pourquoi les célibataires sortent plus que les personnes mariées. On pourra faire à ce propos diverses hypothèses, qu'on pourra toutes formuler plus ou moins facilement en termes de composantes de la situation matrimoniale. On risque évidemment de se laisser ainsi entraîner dans une régression à l'infini, chaque nouvelle variable introduite pouvant à son tour être ramenée à ses diverses composantes. En pratique, le risque n'est pas grand : non seulement on est très vite contraint de s'arrêter, faute d'informations ou faute d'un échantillon de taille suffisante, mais surtout certaines relations apparaissent comme suffisamment évidentes pour qu'on n'éprouve pas le besoin de les spécifier plus.

L'analyse en termes de composantes, qu'on peut effectuer à propos de toutes les variables dont la signification psychologique ou sociologique est complexe, est intéressante également parce qu'elle constitue une possibilité de vérification empirique de ce qu'on appelle d'habitude « interprétation », et qui consiste en général à imaginer un mécanisme reliant les deux variables.

c) On interprète Z comme un *élément intermédiaire de la chaîne causale* qui va de X à Y (ou l'inverse).

On observe une corrélation entre la C.S.P. d'une personne et celle de son père. On sait également qu'il existe une relation entre les études suivies et la profession, et aussi

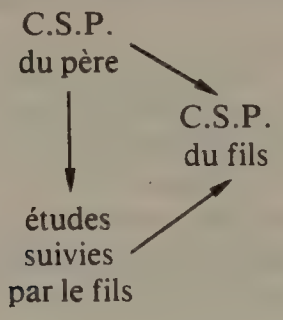
entre ces études et la profession du père. Si l'on veut apprécier correctement le processus de mobilité sociale, il faut se demander si le lien entre la C.S.P. du père et celle de son enfant est due uniquement au fait que celle-ci détermine les études suivies, ou si elle a également un effet spécifique. En d'autres termes, on peut faire l'hypothèse d'une chaîne causale qu'on peut représenter par :

C.S.P. du père → études suivies par le fils → C.S.P. du fils

Ce *schéma de causalité* pourra être considéré comme vérifié si la corrélation entre les C.S.P. du père et du fils disparaît lorsqu'on contrôle les études suivies. Cela indiquera que si le fils d'une personne de C.S.P. basse arrive à suivre des études supérieures, elle atteindra probablement les mêmes C.S.P. que ceux dont le père a une C.S.P. élevée. Inversement, si le fils d'une personne de C.S.P. élevée n'a pas suivi d'études supérieures, elle n'a qu'une faible probabilité d'atteindre une C.S.P. semblable.

Autrement dit, ce schéma traduirait le fait que ce sont les études suivies qui sont déterminantes, la C.S.P. du père n'intervenant que pour faciliter ou freiner l'accès, ou le succès, à ces études.

Une autre hypothèse pourrait être représentée par le schéma suivant :



qui indique, en plus, un effet spécifique de la C.S.P. du père, même à niveau d'études égal. Dans ce cas, la corrélation entre les deux C.S.P. se maintiendra, même si l'on contrôle selon les études suivies.

Les schémas de causalité linéaire, du premier type, peuvent se retrouver dans des cas très divers. On peut être amené à faire des hypothèses de cette forme lorsqu'on veut étudier, par exemple, les relations entre une caractéristique individuelle, une attitude et un comportement. Il est fréquent que le schéma suivant se vérifie :

caractéristique → attitude → comportement

Nous avons ainsi proposé trois interprétations différentes de la même observation, celle que la corrélation entre deux variables disparaît dans des groupes homogènes du point de vue d'une troisième, qui était elle-même en corrélation, dans l'échantillon total, avec les deux premières, ou au moins l'une des deux. On ne peut donc pas choisir l'interprétation qu'on adoptera sur des seuls critères empiriques. Une fois de plus, c'est la représentation qu'on se fait du mécanisme représenté qui guidera le choix. La

corrélation entre l'âge et le niveau d'instruction du premier exemple apparaît comme contingente. Si l'on montre que c'est cette dernière variable qui est déterminante, on n'apprend rien de plus en ajoutant qu'elle est liée à l'âge. Dans le deuxième exemple, au contraire, la variable de l'état matrimonial peut nous apparaître comme tellement liée à l'âge que nous pourrons continuer à parler de lui, même si l'on a précisé la nature de la liaison. Dans le troisième cas, les liaisons sont à priori moins évidentes, il est donc intéressant de conserver un schéma causal complexe qui les fait intervenir toutes les deux, même si l'on parvient à conclure qu'il n'y a pas de lien direct entre X et Y. Toutefois, on voit bien qu'il s'agit de jugements qui restent contestables, et peut-être d'autres chercheurs feraient d'autres interprétations que les nôtres.

Par exemple, dans une enquête sur la mobilité (les déplacements) des personnes âgées, on a constaté qu'une attitude, le *désengagement*, qui exprime le refus de la société, et l'absence d'intérêt qu'on lui porte, présentait une corrélation forte avec le nombre de sorties. Dans la mesure où il s'agissait d'une échelle construite ad hoc, sa valeur explicative restait limitée. Mais les deux variables en question étaient elles-mêmes liées à la catégorie socio-professionnelle avant la retraite, et avec un indicateur d'aliénation dans le travail (fondé sur le caractère répétitif des tâches et la possibilité ou non d'organiser soi-même son travail). En croisant ces quatres variables, on aboutit, en simplifiant, au schéma suivant :

C.S.P. → aliénation dans le travail → désengagement → mobilité

Les trois flèches n'ont pas exactement la même signification. L'aliénation dans le travail peut être considérée comme une composante de la C.S.P., et non comme un des maillons de la chaîne causale, alors qu'on peut dire que la C.S.P. (à cause de l'aliénation qui lui est liée) entraîne le désangagement, qui, lui-même, détermine la mobilité. Mais on peut aussi estimer que la classe sociale, dont la C.S.P. est un indicateur, est fondamentale et que, seule, la relation entre C.S.P. et mobilité est vraiment explicative, les deux autres variables ne faisant que décrire un mécanisme intermédiaire.

Passons maintenant aux cas où la corrélation entre X et Y ne disparaît pas lorsqu'on contrôle Z. Si la relation entre X et Y reste absolument identique dans chaque modalité de Z, la conclusion est simple : Z n'est pour rien dans la relation en question. Mais on peut aussi obtenir d'autres relations, comme celle de la figure 13.

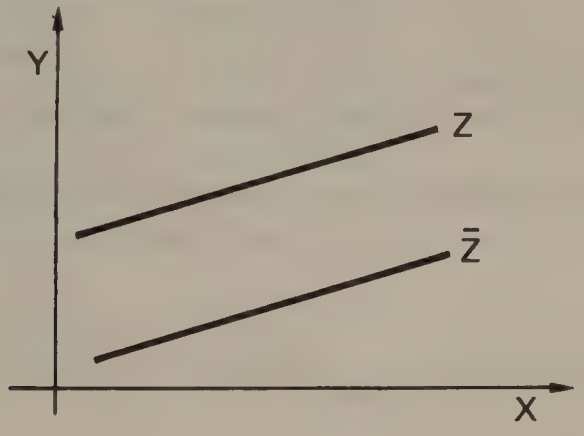
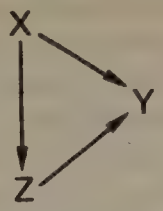


Fig. 13

Nous en avons vu un exemple ci-dessus, à propos des relations possibles entre les C.S.P. du père et du fils. Le schéma correspondant à la figure 13 est le suivant :



Les différences entre les valeurs moyennes de YZ et $Y\bar{Z}$ ne dépendent pas de X. L'effet de Z vient simplement s'ajouter à celui de X.

Mais on pourrait aussi observer le diagramme de la figure 14.

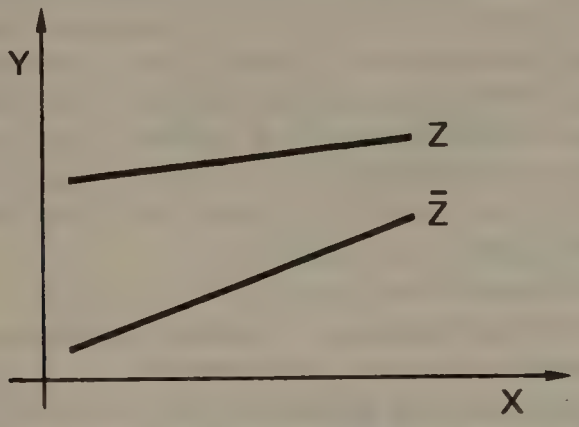
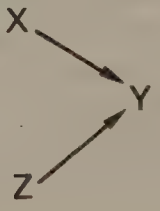


Fig. 14

Les pentes des deux droites sont très différentes : l'« aide » apportée par la C.S.P. élevée du père est nettement plus forte chez ceux qui n'ont pas fait d'études supérieures que chez ceux qui en ont fait. Nous avons là un exemple d'une *interaction*, dont nous avons parlé au chapitre précédent. Le schéma de causalité est moins commode à représenter ; on utilise souvent la notation suivante :

La flèche supplémentaire indique quelle relation est modifiée par l'intervention de la troisième variable.

Avec trois variables, d'autres schémas sont possibles ; par exemple :



indique que X et Z sont indépendants (une corrélation apparente peut intervenir à cause de la liaison avec Y, mais on la fait disparaître en contrôlant celle-ci) et ils exercent tous deux une influence sur Y.

7.3.4. Les relations entre plus de trois variables

7.3.4.1. L'ANALYSE DE CAUSALITÉ

Dans ce qui précède, nous n'avons fait que constituer des tableaux croisant trois variables ou, ce qui revient au même, des séries de tableaux à double entrée, et nous les avons simplement lus. Nous verrons plus loin (voir 7.3.5.) qu'on peut s'aider pour cela de techniques de statistique descriptive, comme des coefficients d'association ou de corrélation, et de tests statistiques. Mais l'essentiel reste la lecture directe des tableaux. Nous pensons que, tant que cela reste possible, ce contact immédiat avec les données est préférable à l'usage de techniques plus raffinées, parfois plus puissantes, mais qui risquent, en ramenant les résultats à quelques coefficients, de constituer un écran entre eux et le chercheur. Même lorsqu'on juge utile ou nécessaire d'utiliser ces méthodes plus systématiques, cela ne devrait pas dispenser d'une lecture plus simple.

Toutefois, lorsqu'on veut pousser l'analyse assez loin, on ne peut pas toujours se limiter à des relations entre trois variables ; la simple lecture des tableaux devient alors pratiquement impossible : on est très vite dépassé par la complexité des relations à comparer [1], dont le nombre croît très rapidement. Différentes techniques ont été proposées, qui ont en commun le pouvoir de prendre simultanément en considération un grand nombre de variables : analyse de causalité, segmentation, analyse factorielle, analyse typologique. Toutes ces techniques ont été rendues pratiquement utilisables par la possibilité de recourir à des ordinateurs : à la main, elles deviennent vite impraticables. Nous n'en présenterons ici que les traits généraux, renvoyant chaque fois à des ouvrages spécialisés pour le détail des techniques. Il existe d'ailleurs des programmes d'ordinateur disponibles.

Parmi ces techniques complexes, celles qui correspondent le mieux aux problèmes que nous avons posés au cours des chapitres précédents sont celles qu'on regroupe sous le nom d'*analyse de causalité*, ou plus prudemment d'*analyse de dépendance*. Elles permettent de substituer à des explorations relativement informelles l'usage d'instruments mathématiques débouchant sur la vérification d'hypothèses causales complexes.

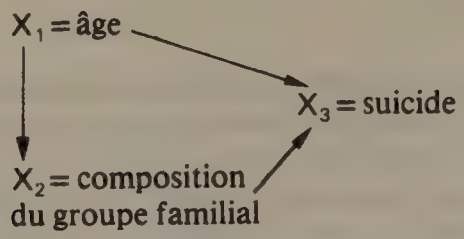
L'idée de départ est due à H.-A. Simon, dont les publications sur ce sujet datent du début des années cinquante. Elle a ensuite été reprise, développée et rendue opérationnelle, principalement par Blalock, puis par Boudon.

L'idée de base des modèles d'analyse de la causalité consiste à représenter un schéma de causalité par un *système récursif d'équations linéaires* [2] exprimant les relations sup-

1. C'est pourtant ce que Durkheim a fait, sans recourir à aucune technique statistique, dans son étude sur le suicide. Mais certaines de ces conclusions s'en ressentent.

2. Un système d'équation est dit récursif si elles sont ordonnées, les valeurs tirées de la première équation étant introduites dans la deuxième, et ainsi de suite. C'est une manière d'exprimer mathématiquement certaines des asymétries du schéma de causalité.

posées par le schéma. Pour donner une idée du fonctionnement de la méthode, nous emprunterons un exemple à Boudon, qui formalise un des aspects de la théorie du suicide de Durkheim. La variable à expliquer est le taux de suicide, et les deux variables retenues sont l'âge et la composition du groupe familial. L'hypothèse est que ces deux facteurs agissent directement sur le suicide, mais que, en plus, l'âge influence la composition du groupe familial.



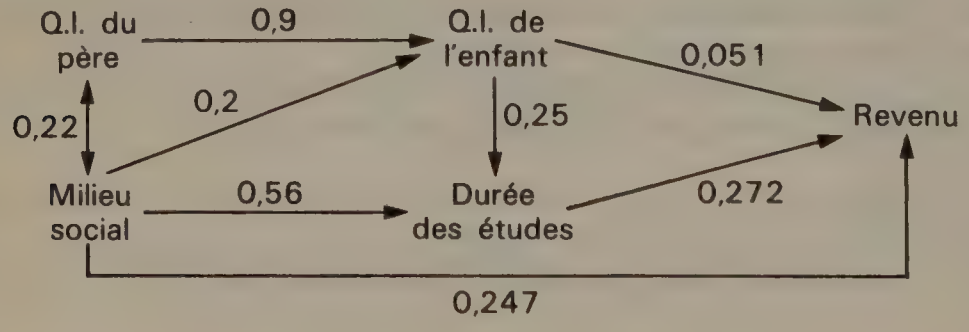
On peut exprimer ces relations par les trois équations linéaires suivantes :

$$x_1 = e_1 \quad (1)$$
$$x_2 = a_{12}\, x_1 \pm e_2 \quad (2)$$
$$x_3 = a_{13}\, x_1 \pm a_{23}\, x_2 \pm e_3 \quad (3)$$

Les e_1, e_2, e_3, représentent des variables aléatoires de moyennes 0, et peuvent être considérées comme des « erreurs », autrement dit comme des sources de variation non explicitement incluses dans le système causal. L'équation (1) indique que l'âge est une donnée exogène, l'équation (2) que la composition de la famille est fonction de l'âge, et d'autres variables non explicitées, et finalement l'équation (3) exprime que le taux de suicide est fonction des deux variables précédentes.

La résolution de ce système nous donne les valeurs des coefficients a_{ij}, qui peuvent être interprétés comme des mesures de l'influence causale de chaque facteur. On peut également évaluer l'ajustement du modèle aux données, c'est-à-dire la part de la variance expliquée par lui, et apprécier si l'ensemble des hypothèses qu'il exprime peut être considéré comme vérifié.

Ce type de schéma de causalité peut être généralisé à un nombre plus grand de variables ; nous avons vu que c'est dans ces cas que la méthode devient précieuse, puisque l'analyse directe des tableaux le révèle impossible. En voici un exemple, indiquant les relations entre le revenu d'une personne et différents facteurs explicatifs



Les valeurs indiquées à côté des différentes flèches sont celles des coefficients des équations. Elles estiment donc la pente des droites représentant les relations entre variables. On voit que le Q.I. de la personne considérée affecte très peu directement son revenu, mais plus fortement par l'intermédiaire de la durée des études. Le « milieu social » apparaît comme un facteur essentiel, puisqu'il intervient à la fois directement sur le revenu, et indirectement par l'intermédiaire du Q.I. et de la durée des études.

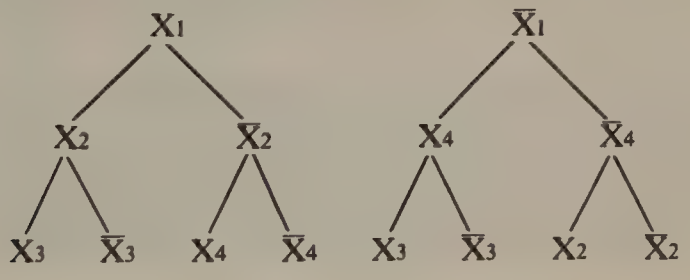
Ces exemples ne font pas intervenir d'*interaction* entre les variables. L'introduction de celle-ci complique non seulement le schéma causal lui-même, mais aussi l'aspect strictement mathématique, puisqu'on ne peut plus se contenter de relations linéaires. Mais, dans la plupart des cas, des méthodes de résolution existent.

7.3.4.2. La méthode de segmentation

Dans ce qui précède, nous avons toujours supposé que le chercheur avait des hypothèses plus ou moins claires sur les relations qui peuvent exister entre les variables qu'il a recueillies. C'est une vue un peu idéalisée de la pratique des enquêtes ; certes, introduire des questions relatives à une variable traduit bien l'idée plus ou moins vague que celle-ci peut jouer un rôle, mais l'hypothèse ne va souvent pas plus loin. Dans ce cas, plutôt que d'effectuer systématiquement tous les croisements de deux, trois variables, ou plus, en recherchant ensuite où se manifestent des relations, il sera intéressant de disposer d'une méthode plus économique, permettant de mettre en évidence rapidement des combinaisons de variables explicatives. La méthode dite de *segmentation*, proposée initialement par Belson, répond dans certains cas à ce problème.

La méthode s'applique lorsqu'on cherche à expliquer une variable par un certain nombre d'autres, toutes ces variables étant dichotomiques, ou pouvant être sans trop de perte regroupées pour se ramener à des dichotomies. Par exemple, une des premières applications de Belson a été d'essayer d'expliquer la lecture d'un journal par différentes caractéristiques.

On calcule la corrélation entre la variable à expliquer et chacune des caractéristiques, et on retient celle qui donne la corrélation la plus élevée, X_1. On divise donc l'échantillon en deux sous-échantillons, correspondant à X_1 et $\overline{X}_1$; par construction, ils peuvent être très différents selon la variable à expliquer. On répète l'opération séparément dans chaque sous-échantillon, retenant dans chaque cas la caractéristique qui a la meilleure corrélation avec la variable à expliquer, caractéristique qui peut ne pas être la même dans chaque sous-échantillon, et on continue, aboutissant à la fin à une structure d'arbre :



Ce schéma indique que c'est X_1 qui est la variable la plus discriminative sur l'ensemble de l' échantillon. Dans le groupe homogène caractérisé par X_1, c'est X_2 qui est discriminant, alors que dans le groupe $\overline{X}$ c'est X_4, etc. On forme ainsi des groupes, dans l'arbre ci-dessus, définis par $X_1\ X_2\ X_3$, $X_1\ X_2\ \overline{X}_3$, $\overline{X}_1\ X_2\ X_4$, etc., constitués de telle manière que leur différence du point de vue de la variable à expliquer soit aussi grande que possible. Comme toutes les méthodes qui ne procèdent pas par vérification d'hypothèses, la segmentation [1] donne parfois des résultats difficiles à interpréter, mais elle peut avoir un grand intérêt heuristique.

7.3.4.3. L'ANALYSE DES STRUCTURES DE RÉPONSES

Toutes les analyses qui précèdent portaient sur des relations asymétriques, avec éventuellement la possibilité de supposer des rétroactions. Lorsque toutes les variables peuvent être considérées comme de même nature, jouent des rôles équivalents dans les interprétations, nous avons vu que les relations entre elles sont interprétées comme symétriques. Les schémas de causalité du paragraphe précédent ne s'appliquent donc plus, même sous une forme simple. Cela ne veut évidemment pas dire que l'analyse des relations entre de telles variables soit sans intérêt, en particulier celle des relations entre plus de deux variables. Mais le mode d'approche le plus intéressant est tout différent de celui de la causalité, même prise dans son sens le plus large.

Partons du cas, très banal, d'une enquête d'opinion comportant plusieurs questions. Supposons qu'on ait posé une série de questions portant sur différents aspects de la politique gouvernementale. Lorsqu'on publiera les résultats, on indiquera en général les pourcentages de réponses favorables (et éventuellement, en plus petits caractères, ceux des réponses défavorables et des non-réponses) à chaque question. Si plusieurs de ces pourcentages dépassent 50 %, on en concluera facilement que la majorité de la population approuve l'ensemble de la politique qu'on lui demande de juger. Ou, si deux propositions opposées obtiennent des pourcentages voisins d'approbation, on pourra en retirer l'impression que les sujets ont répondu de façon contradictoire. Il est évident que ces deux conclusions, bien que compatibles avec les résultats communiqués, sont abusives. Pour en vérifier la validité, il faudrait étudier la relation entre les diverses réponses. Ce n'est que si on peut montrer qu'il existe des corrélations élevées entre les réponses aux questions sur les différents aspects de la politique gouvernementale qu'on pourra conclure valablement à une attitude générale d'approbation ou de désapprobation de celle-ci dans son ensemble. Mais il se peut aussi que ce ne soient pas les mêmes sujets qui répondent favorablement aux différentes questions, auquel cas il n'est pas légitime de parler d'une approbation générale. De même, on ne peut qualifier les sujets de contradictoires que si l'on s'est assuré que ce sont bien les mêmes qui ont répondu de façon non cohérente [2].

---

1. Le terme de segmentation provient des études du marché ; la méthode permet, en prenant comme variable à expliquer l'achat d'un produit, de constituer des « segments » de clientèle, avec leurs caractéristiques.
2. Même si c'est le cas, la conclusion reste encore trop rapide. On plaisante souvent sur l'absurdité des Français dont l'idéal, en matière de logement, selon certaines enquêtes, serait un pavillon avec jardin sur les Champs-Elysées. Un tel résultat est normal, si on a posé séparément une question sur les préférences entre logement individuel ou collectif, et une autre sur la localisation souhaitée. Cette décomposition des questions ne permet évidemment pas aux sujets d'exprimer les arbitrages qu'ils sont disposés à faire, et que le chercheur, ou son commentateur, effectueront ensuite trop facilement à leur place, ou oublieront de faire, comme dans l'exemple ci-dessus.

Indépendamment de ces précautions contre des interprétations abusives, il peut être intéressant d'étudier la structure des réponses à un ensemble de questions portant sur des thèmes apparentés. Il s'agit d'étudier les combinaisons de réponses données par les mêmes sujets, et voir s'il existe des relations entre ces combinaisons. La structure hiérarchique (cf. 7.2.1.2.) en constitue un exemple, caractérisé par l'existence d'une relation d'inclusion entre les différents ensembles de réponses, qui permet de conclure à une variable unidimensionnelle sous-jacente unique. L'existence de corrélations entre les réponses prises deux à deux, ou entre des combinaisons de réponses à des sous-ensembles de questions, indique également une variable unidimensionnelle, mais dans un tout autre sens. Nous avons discuté précédemment de la différence entre ces deux approches (cf. 7.3.2.3.), nous n'y reviendrons donc pas ici.

Ces deux exemples sont caractéristiques de deux manières d'aborder le problème de l'analyse des structures de réponses. L'une se fonde sur la prise en considération des relations entre ensembles de réponses. Elle utilise principalement l'algèbre de Boole, et ne fait aucune hypothèse métrique. Le résultat de l'analyse peut être la mise en évidence de relations d'implication entre réponses, d'ordres partiels, etc. L'effort le plus poussé dans cette direction a été celui de Claude Flament. Son « analyse algébrique des questionnaires » vise à dégager une structure élémentaire permettant, au moyen de certaines opérations, d'engendrer l'ensemble des combinaisons de réponses observées, la structure de base pouvant être considérée comme un résumé de l'ensemble des combinaisons de réponses obtenues.

L'autre approche part de l'analyse des corrélations entre réponses, et utilise l'algèbre linéaire, donc suppose une métrique. Outre les méthodes de construction d'échelles d'attitude dont nous avons parlé ci-dessus (cf. 7.2.1.2.), elle a débouché, dans son développement, sur l'analyse factorielle et ses dérivés, qui présentent l'intérêt de ne pas être liée à l'hypothèse d'unidimensionnalité.

7.3.4.4. Remarques sur l'analyse des données

De ce qui a été dit au paragraphe précédent, on peut retirer l'impression que les méthodes habituelles d'analyse des relations perdent des informations importantes, parce qu'elles ne permettent de tenir compte que d'un petit nombre de variables à la fois. Nous avons vu que la prise en considération d'une variable supplémentaire peut modifier complètement les interprétations qu'on peut proposer, permet d'avancer des hypothèses beaucoup plus fines et de mettre en évidence des mécanismes plus complexes. Pourquoi alors ne pas ajouter deux, trois, dix variables ? On améliorerait encore l'analyse. Toutefois, nous l'avons vu, on se heurte très vite à deux limites : l'une liée à notre incapacité à saisir et à interpréter, sans instruments spécifiques, des relations entre plus de trois ou quatre variables ; l'autre due à la taille des échantillons usuels, qui ne permet pas de constituer les tableaux voulus. Dans une certaine mesure, l'analyse de causalité cherche à surmonter ces difficultés. Mais on espère toujours disposer de méthodes globales d'analyse de l'ensemble du questionnaire.

Depuis quelques années, on a vu se développer, et se répandre largement, sous la dénomination globale d'*analyse des données*, un ensemble de méthodes qui, cherchant à traiter simultanément un grand nombre de variables, semblent bien répondre à cette préoccupation. Leur développement rapide, très récent, est dû principalement à la possibilité d'utiliser des ordinateurs, mais les idées directrices en sont connues depuis longtemps. Plus précisément, on peut les faire remonter aux premiers travaux de Spearman sur l'analyse factorielle, en 1904-1905.

Avant de se présenter, comme elle le fait aujourd'hui, comme une méthode neutre du point de vue théorique, et applicable à une gamme de données très large, l'analyse factorielle a été, à ses débuts, l'expression mathématique d'une théorie particulière de l'intelligence, la théorie dite « des deux facteurs », puis, plus largement, d'une théorie des aptitudes. Pendant une cinquantaine d'années, bien que s'étant peu à peu dégagée de ses premières attaches théoriques, elle n'a guère été utilisée que par des psychométriciens, c'est-à-dire par des psychologues travaillant avec des *tests* ; ce sont eux également qui, pendant cette même période, l'ont perfectionnée et développée.

Approximativement à partir de la fin des années cinquante, la possibilité d'utiliser des ordinateurs de plus en plus puissants, en rendant plus facile son emploi, a attiré sur l'analyse factorielle l'attention de chercheurs de disciplines très différentes : biologistes, économistes, géographes, sociologues, etc., qui se trouvent souvent devoir travailler avec des ensembles très vastes de données, et qui ont cherché dans cette méthode, ou dans ses dérivés, un moyen d'y mettre de l'ordre, de les décrire de façon commode, de les classer, de les organiser. En même temps, un nombre croissant de mathématiciens et de statisticiens y ont trouvé des thèmes de recherche, développant des méthodes de plus en plus générales, de moins en moins liées, en principe, à une théorie particulière ou à une classe déterminée d'hypothèses sur la structure des données qu'on veut analyser. C'est sous cette forme générale, et facilement utilisable en pratique, que ces méthodes sont revenues chez les psychologues, et en particulier, c'est ce qui nous intéresse ici, chez les spécialistes des enquêtes qui, dans la majorité, avaient ignoré l'analyse factorielle classique. A première vue, en effet, les différentes formes de l'analyse des données semblent particulièrement bien adaptées au traitement des résultats d'enquêtes, qui consistent à réunir, sur chaque sujet, un grand nombre d'informations. Nous avons vu plus haut (7.1.1.) que ces résultats peuvent, en principe, toujours se mettre sous la forme d'une matrice rectangulaire, avec les sujets en ligne et les questions en colonne, ou l'inverse. Les méthodes d'analyse dont nous avons parlé au paragraphe précédent reviennent toutes soit à isoler un petit nombre de colonnes, et à étudier les relations entre elles, soit à essayer de constituer des groupes de lignes suffisamment semblables. Comme nous l'avons souligné, on perd, ce faisant, une certaine information, puisqu'on fractionne les résultats, on ne les saisit plus dans leur totalité. *C'est précisément cette saisie globale que vise l'analyse des données.*

Une autre raison de l'attrait de cet ensemble de méthodes, c'est qu'elles sont *multidimensionnelles*, contrairement aux techniques de construction d'indicateurs que nous avons exposées plus haut et qui, elles, ne s'appliquent qu'à des données unidimensionnelles, et donc laissent leur utilisateur impuissant lorsque cette hypothèse n'est pas vérifiée. Certes, disposer de variables unidimensionnelles (c'est-à-dire qui prennent

leurs valeurs dans un ensemble totalement ordonné) est très commode si l'on veut faire des comparaisons, calculer des corrélations, etc. Mais les données ne s'y prêtent pas toujours ; ces méthodes unidimensionnelles constituent donc des *modèles* exprimant des hypothèses sur la structure des données, hypothèses qu'on pourra accepter ou rejeter. En revanche, l'analyse des données s'applique en principe à toute donnée, ce qui constitue à la fois sa puissance et sa faiblesse. Sa puissance, parce qu'on ne se trouve jamais devant un cas où elle ne s'applique pas [1] ; sa limite, parce que n'étant guidé par aucune hypothèse on se trouve souvent très embarrassé devant les résultats obtenus, et incapable de les interpréter.

Le fait que l'analyse des données ne vérifie pas d'hypothèses précises est souvent présenté comme une preuve de sa neutralité théorique. Mais il ne faudrait pas en conclure qu'il n'y a pas, à la base, de postulats non vérifiables qu'on applique aux données, et donc aux résultats qu'on en tire. C'est là que des présupposés théoriques peuvent se glisser, mais il est très difficile d'en apprécier la signification et les conséquences.

Pas plus que nous ne l'avons fait pour l'analyse de causalité, il n'est dans notre propos, ici, d'entrer, même de façon sommaire, dans le détail des différentes techniques proposées, dont l'exposé exigerait à lui seul un gros volume, et qui serait d'ailleurs en dehors de notre compétence. Nous nous bornerons à en exposer les principes généraux, et à évaluer leur apport dans le domaine des enquêtes.

L'idée générale, commune à toutes les méthodes d'analyse des données, est la suivante. Chaque unité (chaque sujet, dans le cas des enquêtes) est caractérisé par la suite de ses valeurs selon les n variables prises en considération (par exemple ses réponses aux n questions posées). On peut donc la représenter par un point dans un espace à n dimensions. Mais, en général, les points représentant les différents sujets ne sont pas répartis uniformément dans cet espace [2] : ils sont regroupés dans certaines zones, ou selon certaines directions. Il est alors possible de décrire ce même ensemble de points au moyen d'un plus petit nombre de dimensions. Par exemple si, dans un plan, tous les points se trouvent alignés selon une droite, on pourra caractériser chaque point par une seule valeur correspondant à sa position sur cette droite, au lieu de ses deux coordonnées sur les deux axes qui définissent le plan. Si, sans être rigoureusement alignés, les points sont néanmoins tous proches d'une même droite, prendre celle-ci comme axe unique pourra encore constituer une approximation acceptable. On peut donc décrire des données bi-dimensionnelles par une seule dimension, à condition que les deux dimensions initiales soient en corrélation.

En termes généraux, l'objectif de l'analyse factorielle est de chercher le nombre minimum de dimensions indépendantes (les facteurs) capables de rendre compte, avec une perte d'information faible, d'un ensemble de données décrites par un plus grand nombre de variables corrélées entre elles. On dégage ainsi ce qu'il peut y avoir de commun à

---

1. Cette affirmation est trop absolue. Mais les limites de l'analyse de données, dont nous ne parlerons pas ici, sont très larges, et en pratique on les rencontre rarement. En revanche, il arrive souvent que les résultats ne soient pas jugés « intéressants » par le chercheur, ce qui est une autre manière de rejeter la structure sous-jacente.
2. Si c'était le cas, cela indiquerait que toutes les variables sont indépendantes ; toute analyse portant sur les relations entre variables, ou entre sujets, deviendrait donc inutile.

plusieurs de ces variables, on explicite des variables nouvelles susceptibles de rendre compte de leur concomitance.

Ainsi, pour prendre un exemple dans le domaine des attitudes politiques, Eysenck a montré qu'on pouvait ramener un grand nombre de positions politiques sur des points divers à deux grandes dimensions : l'une qu'on peut interpréter comme l'opposition gauche-droite classique, et une autre moins bien connue, dont il désigne les deux pôles par les expression *tough-minded* et *tender-minded*, ce qu'on peut à la rigueur traduire, mal, par les « durs » et les « mous ».

Chez Spearman, le recours à l'analyse factorielle était l'expression d'une position théorique, qui supposait l'existence d'une « intelligence générale » unique ; il s'agissait donc d'une hypothèse particulière sur les résultats de l'analyse, puisque, selon l'hypothèse, on ne devait trouver qu'un seul facteur général rendant compte de toutes les corrélations qu'on peut observer entre les tests dits d'intelligence. Mais, sous sa forme moderne, sans hypothèses sur le nombre et la nature des facteurs, la fonction principale de l'analyse factorielle est de mettre en ordre les données, de les résumer sous une forme qui, en principe, doit être plus claire. Sous une forme plus complexe, c'est la même démarche que lorsqu'on réduit un ensemble de valeurs à leur moyenne. C'est en cela qu'elle apparaît comme neutre du point de vue théorique : l'analyse des données, dans l'idée de ceux qui l'ont développée, serait strictement descriptive. Cette neutralité a d'ailleurs été contestée ; en effet, toutes les techniques utilisées, quelles que soient leurs spécificités, s'appuyent sur des postulats souvent implicites (au moins pour le psychologue utilisateur), par exemple dans leur manière de définir la distance entre points dans l'espace considéré. Il est toutefois difficile de se rendre compte des conséquences de ces postulats, et de voir quelle est leur traduction en termes de présupposés psychologiques ou sociologiques.

Le caractère descriptif de l'analyse factorielle a pour conséquence qu'elle ne peut, le plus souvent, constituer qu'une étape dans le processus de recherche. Si, devant un ensemble de données à la fois nombreuses et complexes, il est souvent très utile de constater que les variables peuvent être réduites à un petit nombre, ou de classer les unités étudiées en types, cela ne dispense pas de la vérification d'hypothèses précises, vérification qui constitue une des caractéristiques principales de l'activité scientifique.

L'analyse de données traite toutes les relations comme symétriques. Or, nous l'avons vu, il est difficile, et il n'est pas souhaitable, d'éviter de les interpréter en termes de causalité, de distinction entre variables explicatives et expliquées, d'influence d'une variable sur une autre, etc., de telles interprétations étant, elles, asymétriques. Appliquée inconsidérément à des données hétérogènes, entre lesquelles on peut supposer des relations interprétables de façons très diverses, l'analyse des données les traitera toutes de la même manière, ce qui, au lieu d'éclairer, peut rendre les données encore plus confuses.

A notre sens, l'analyse factorielle trouve sa plus grande justification, et se révèle particulièrement utile, dans les cas où l'on dispose d'un grand nombre de variables *de même nature*, entre lesquelles on est disposé à interpréter toutes les relations de la même manière, et de façon symétrique. Par exemple, lorsqu'on a posé un grand nombre de questions d'attitude, une analyse factorielle permettra de dégager les dimensions sous-

jacentes à cet ensemble ; les corrélations observables entre deux questions sont alors interprétées comme l'expression d'une même attitude sous-jacente, l'analyse factorielle permettant d'estimer celle-ci. Mais si l'on avait traité simultanément ces mêmes questions et aussi des données socio-économiques sur les individus, l'interprétation des facteurs extraits aurait été beaucoup plus hasardeuse.

A côté de l'analyse factorielle, l'autre grand groupe de méthodes d'analyse des donnés vise à construire des typologies (ou taxinomies). Alors que l'analyse factorielle s'intéresse aux *variables*, les méthodes de typologie portent sur les unités étudiées, les sujets d'une enquête par exemple [1]. L'objectif est de constituer des groupes d'unités semblables, telles que celles qui font partie d'un même groupe soient plus proches les unes des autres qu'elles ne le sont des unités des autres groupes.

Remarquons que cette définition de la typologie, qui suppose une structure très particulière des données, n'est pas la seule possible. Elle est différente, par exemple, de celle que propose Lazarsfeld. Celui-ci définit ses groupes en combinant également plusieurs variables, mais sans se préoccuper de la distribution des sujets dans l'espace ainsi défini. Il définit plutôt ses groupes soit à priori, soit en fonction d'un critère extérieur.

Les types définis par l'analyse des données ne doivent pas non plus être confondus avec les *types idéaux* de Max Weber, constructions théoriques permettant d'analyser la réalité, mais dont on ne s'attend pas à ce qu'ils puissent se réaliser empiriquement.

Toutes ces définitions ont en commun de chercher à classer des individus *qualitativement*, terme assez vague mais qui, *en général*, signifie « *autrement qu'en les ordonnant totalement selon une seule variable* », sauf peut-être chez Max Weber, dont le type idéal n'a pas la même fonction classificatrice. En combinant plusieurs variables même ordinales ou mesurables, on définit des classes qui ne sont plus totalement ordonnées. Les méthodes de construction de typologies dérivées de l'analyse des données supposent, en plus, que seules certaines combinaisons des valeurs des variables initiales sont représentées.

Comme pour l'analyse factorielle, construire des typologies à partir de variables très hétérogènes nous semble dangereux, car on court le risque de masquer, dans la définition même des types, des relations qui pourraient être intéressantes en elles-mêmes. Par exemple, plutôt que de définir les types à la fois par des opinions, des comportements et des caractéristiques socio-économiques, il peut être plus intéressant de définir une première typologie à partir des seules opinions, une autre à partir des comportements, et ensuite d'étudier les relations entre ces deux typologies, et entre chacune d'elles et les variables socio-économiques, ce qui permet de vérifier plus explicitement des hypothèses sur les relations entre ces catégories de variables.

Il est intéressant de rapprocher de l'analyse des données, méthode mathématique, une autre technique, qui poursuit les mêmes objectifs avec des moyens très différents. Il s'agit des *matrices* imaginées par J. Bertin. C'est une matérialisation de la matrice globale dont nous avons parlé plusieurs fois (cf. 6.6. et 7.1.1.), chaque case étant représentée par un cube, la « valeur » correspondante se marquant, par du noir, du

1. Dans les faits, il n'y a pas lieu d'opposer analyse factorielle et recherche d'une typologie qui se fait, selon certaines méthodes, après réduction de l'espace des données, donc après analyse factorielle.

blanc, ou différentes plages de gris. Un système ingénieux de tiges permet d'embrocher toute une ligne, ou toute une colonne, et donc de les permuter de manière à regrouper les lignes ou les colonnes semblables, en se guidant sur les caractéristiques visuelles des figures ainsi formées. On peut ainsi regrouper soit des sujets, ce qui revient à constituer une typologie, soit des variables, mettant ainsi en évidence des groupes de facteurs.

Toutes ces méthodes d'analyse des données ont en commun de ne pas être des techniques de vérification d'hypothèses préalables, mais de rester descriptives. C'est à la fois leur intérêt et leur limite. Elles permettent de classer, elles indiquent des proximités, des ressemblances, mais elles ne mettent pas en évidence l'existence de lois. Le plus souvent, elles ne peuvent donc constituer qu'une étape intermédiaire dans une recherche.

7.3.5. La lecture et la présentation des tableaux croisés

Mis à part les cas d'utilisation de techniques globales complexes, comme l'analyse de causalité ou l'analyse factorielle, l'essentiel des résultats d'enquêtes se présente sous la forme de tableaux à double entrée, ou *tableaux croisés*. Toutes les analyses dont nous avons parlé au cours des chapitres précédents portent sur de tels tableaux, et éventuellement, lors de l'analyse de relations entre plus de deux variables, sur la confrontation de plusieurs d'entre eux. Jusqu'à présent, nous étant intéressés à l'*interprétation* des relations qui s'expriment par ces tableaux, nous n'avons guère parlé de la manière de reconnaître ces relations, autrement dit de leur lecture. Maintenant, ayant à l'esprit ce qu'on peut y chercher, nous pouvons examiner la manière de les lire et de les présenter.

Un tableau croisant les variables X et Y se présente sous la forme suivante :

	X_1	X_2	X_i	.....	X_m
Y_1	f_{11}		fi_1		$n\ (Y_1) = \Sigma_i fi_1$
Y_2					
Y_j	f_1j		fij		$n\ (Y_j) = \Sigma_i fij$
Y_n					
	$n(X_1)$		$n(X_i)$		$N = \Sigma_i \Sigma_j fij$
	$\Sigma_j f_1j$		$\Sigma_j fij$		

où $X_1 .... X_m$ représentent les modalités de la variable X,
$Y_1 .... Y_n$ représentent les modalités de la variable Y,
fij le nombre d'observations présentant les modalités X_i et Y_j,
$n(Y_i)$ le nombre d'observations présentant la modalité Y_i.

La distribution des n(X_i) est la *distribution marginale* de X ; c'est ce qu'on obtient par un tri à plat (cf. 7.2.3.).

Si toutes les lignes (ou toutes les colonnes) étaient identiques, c'est-à-dire si toutes les distributions partielles de X pour chaque valeur de Y étaient les mêmes, la conclusion serait immédiate : les deux variables sont indépendantes. Mais, en pratique, cette situation est très rare, puisque les fréquences marginales des différentes modalités ne sont pas les mêmes. On ne peut donc pas comparer directement des fréquences sous peine de commettre des erreurs graves. Par exemple, dans le tableau suivant , on ne peut pas

	X_1	X_2	X_3	
Y_1	30	45	60	135
Y_2	20	30	40	90
	50	75	100	225

commettre des erreurs graves. Par exemple, dans le tableau suivant, on ne peut pas parce qu'il y a plus de X_1Y_1 (30) que de X_1Y_2 (20). Cette différence se retrouve dans les autres modalités de X, et donc dans l'échantillon total. S'il y a moins de X_1Y_2 que de X_1Y_1, c'est simplement parce que, en général, il y a moins de Y_2 que de Y_1. En fait, ce tableau permet de conclure à l'indépendance de X et de Y, puisqu'on retrouve la même proportion de Y_1 et de Y_2 pour chaque valeur de X. On le verrait facilement, même avec des données plus complexes, en calculant les pourcentages correspondants : les

	X_1	X_2	X_3	
Y_1	60 %	60 %	60 %	60 %
Y_2	40 %	40 %	40 %	40 %
	50=100 %	75=100 %	100=100 %	225=100 %

pourcentages ont été calculés par rapport à la somme de chaque colonne (*pourcentages verticaux*), c'est-à-dire par rapport à la fréquence de chaque modalité de X. Mais nous aurions pu aussi bien calculer les pourcentages par rapport aux fréquences des modalités de Y, ce qui aurait donné le tableau suivant (*pourcentages horizontaux*) :

	X_1	X_2	X_3	
Y_1	22,2 %	33,3 %	44,4 %	135 = 100 %
Y_2	22,2 %	33,3 %	44,4 %	90 = 100 %
	22,2 %	33,3 %	44,4 %	225 = 100 %

Dans chaque tableau, l'égalité des pourcentages, en ligne dans le premier, en colonne dans le second, permet de conclure à l'indépendance des deux variables. Avant de nous poser la question de savoir lequel des deux tableaux choisir, il est important de remarquer qu'ils permettent les mêmes conclusions, quant à l'existence d'une relation.

Dans chaque cas, ce qui nous permet de conclure à l'indépendance de X et de Y, ce sont deux propriétés complémentaires, formellement équivalentes :

Si on a calculé les pourcentages « verticaux », (a) *les pourcentages d'une même ligne sont égaux*, et (b) *les pourcentages dans chaque colonne varient de la même manière*. Si ce sont les pourcentages horizontaux qu'on a calculés, c'est l'inverse.

Pour simplifier, on énoncera la règle suivante : *si l'on a calculé les pourcentages par rapport à la somme de chaque ligne (pourcentages horizontaux), on comparera les pourcentages d'une même colonne ; si l'on a calculé les pourcentages par rapport à la somme de chaque colonne (pourcentages verticaux), on les comparera dans une même ligne.*

Si ces pourcentages ne sont pas égaux, ou plutôt s'ils s'écartent trop de l'égalité, on rejettera l'hypothèse d'indépendance, et on conclura donc à l'existence d'une relation. Remarquons que cela ne précise pas la forme que peut prendre cette relation, qui demande une inspection plus fine du tableau. Nous verrons plus loin des techniques statistiques qui fournissent des critères pour accepter ou rejeter l'hypothèse d'indépendance. Toutefois, à notre avis, le recours à ces techniques ne devrait pas dispenser d'un premier examen, « visuel », des tableaux, qui assure au chercheur un meilleur contact avec les données.

Si tout ce qu'on cherche, c'est aboutir à la conclusion d'indépendance des deux variables, ou d'existence d'une relation entre elles, tirer les pourcentages horizontalement ou verticalement est tout à fait indifférent. En revanche, si l'on veut interpréter plus finement la relation, la préciser, l'une des présentations peut se révéler plus commode que l'autre. Cela dépend des rôles que l'on attribue à chacune des deux variables.

Supposons que X soit la catégorie sociale, et Y la présence ou l'absence d'un certain comportement. Reprenant une terminologie issue des études de marché, nous distinguerons :

— la *composition* des sous-populations présentant respectivement Y_1 et Y_2,

— la *pénétration* de Y_1 ou de Y_2 dans chacune des catégories X_1, X_2, X_3.

La composition intervient dans des proportions de la forme : sur 100 personnes présentant le comportement Y_1, p_1 % appartiennent à X_1, p_2 % à X_2, p_3 % à X_3. Elle correspond, dans les tableaux ci-dessus, aux pourcentages horizontaux.

La signification de la pénétration s'exprime par « sur 100 personnes appartenant à X_1, p % présentent le comportement Y_1 ». Elle est donnée par les pourcentages verticaux.

Ces deux notions décrivent évidemment la même situation : si la pénétration de Y_1 dans les différentes catégories n'est pas la même, les compositions de Y_1 et de Y_2 seront nécessairement différentes. Cela correspond à l'existence d'une relation entre X et Y.

Toutefois, malgré leur équivalence formelle, composition et pénétration donnent des renseignements différents, ne mettent pas en évidence les mêmes aspects du phénomène

étudié. Reprenons l'exemple (7.2.3.) de la lecture d'un journal dans trois catégories sociales ; le tableau que nous avions donné indiquait la composition des lecteurs et des non-lecteurs :

	A	B	C	
Lecteurs	60 %	30 %	10 %	200 = 100 %
Non-lecteurs	72 %	24 %	4 %	1000 = 100 %
Total	70 %	25 %	5 %	1200 = 100 %

La comparaison des pourcentages de chaque colonne (puisqu'il s'agit de pourcentages verticaux) montre que la lecture du journal en question est liée à la classe sociale (sous réserve que les différences puissent être considérées comme significatives, ce que nous admettons ici). L'examen de la première ligne (compositions des lecteurs) nous indique que la majorité d'entre eux appartiennent à la catégorie A. La composition avec les deuxième et troisième lignes montre qu'il y a encore plus de A dans la population totale et parmi les non-lecteurs, ce qui pose un problème, mais constater que la majorité des lecteurs appartient à A reste une constatation vraie, et intéressante.

Calculons maintenant les pénétrations :

	A	B	C	
Lecteurs	14 %	20 %	33 %	17 %
Non-lecteurs	86 %	80 %	67 %	83 %
	840=100 %	300=100 %	60=100 %	1200=100 %

L'impression qu'on peut en retirer est complètement modifiée : comme on a calculé les pourcentages séparément dans chaque catégorie sociale, le fait que les A soient fortement majoritaires dans l'échantillon ne se voit plus. Tout ce qui se remarque, c'est qu'il y a une faible proportion de lecteurs en A, et une plus forte en C.

De même que la seule considération des compositions risquait de masquer le succès relatif du journal en question dans la catégorie C, de même, maintenant, on risque d'oublier que, malgré tout, la majorité des lecteurs sont des A.

On voit que le calcul de pourcentages, tout simple qu'il soit, constitue un élément essentiel pour la lecture des tableaux croisés, aussi bien en vue de leur analyse par le chercheur que de la communication des résultats.

Leur utilisation demande toutefois quelques précautions. Calculer des pourcentages, c'est ramener un certain nombre de valeurs à 100, c'est donc en augmenter certaines et en diminuer d'autres. N'indiquer qu'un pourcentage, et non la valeur de la fréquence correspondante, c'est se priver de la possibilité d'apprécier la précision des résultats, et donc de les interpréter en connaissance de cause. Et cela est à la fois vrai du lecteur qui prend connaissance du tableau que du chercheur qui l'interprète et qui, éventuellement, aura recours à des tests statistiques.

On ne jugera en effet pas de la même manière un pourcentage déterminé selon qu'il aura été obtenu sur un échantillon de 40, de 400, ou de 4000 personnes. (Rappelons que, dans le calcul de l'intervalle de confiance d'une proportion, la taille de l'échantillon intervient, en plus de la valeur de cette proportion). On devrait donc toujours indiquer, en plus du pourcentage, les valeurs absolues correspondantes, ou au moins celle de la base, c'est-à-dire la valeur qui correspond à 100 %.

C'est ce que nous avons fait dans les tableaux ci-dessus, nous nous sommes contentés d'indiquer la valeur des différentes bases, pour ne pas surcharger le tableau, et le rendre moins lisible. Mais, et c'est cela qui est essentiel, toute l'information a été communiquée : le lecteur désireux de connaître des fréquences non indiquées a tous les éléments nécessaires pour les calculer.

Il est également indispensable de toujours indiquer dans quel sens sont tirés les pourcentages. Un tableau comme celui-ci ne devrait jamais être présenté, même si, en examinant les pourcentages, le lecteur peut voir que c'est verticalement que leur somme donne 100 :

	X_1	X_2	X_3
Y_1	14 %	20 %	33 %
Y_2	86 %	80 %	67 %

Cela impose au lecteur un effort inutile et augmente les risques d'erreur. D'autre part, il peut arriver que toutes les catégories ne soient pas représentées dans le tableau (la somme sera inférieure à 100 %), ou que les catégories ne soient pas mutuellement exclusives (la somme sera supérieure à 100 %), ce qui complique encore la lecture, si ces précisions ne sont pas indiquées.

Il est aussi nécessaire d'indiquer la signification de la base. S'agit-il, par exemple, de la totalité du sous-échantillon, ou seulement de ceux qui ont donné une réponse ? Nous avons vu plus haut les problèmes de principe que ce choix pose (cf. 2.1.3. et 2.1.4.). Quelle que soit la décision prise, il est indispensable qu'elle apparaisse immédiatement au lecteur.

La présentation de tableaux à trois dimensions est évidemment plus difficile. On peut juxtaposer des tableaux à double entrée, selon la présentation suivante :

		Z				$\bar{Z}$	
	X	$\bar{X}$			X	$\bar{X}$	
Y	fxyz	$f\bar{x}yz$	$fy\bar{z}$	Y	$fxy\bar{z}$	$f\bar{x}y\bar{z}$	$fy\bar{z}$
$\bar{Y}$	$fx\bar{y}z$	$f\bar{x}\bar{y}z$	$f\bar{y}z$	$\bar{Y}$	$fx\bar{y}\bar{z}$	$f\bar{x}\bar{y}\bar{z}$	$f\bar{y}\bar{z}$
	fxz	$f\bar{x}z$			$fx\bar{z}$	$f\bar{x}\bar{z}$	

Lorsque chaque variable, et surtout Z, a un grand nombre de valeurs, cette présentation devient peu commode. On peut recourir aux diagrammes que nous avons souvent utilisés pour notre exposé, en construisant des courbes moyennes de Y en fonction de X pour chaque valeur de Z, ce qui permet des comparaisons plus faciles.

On voit qu'on ne peut pas maintenir une symétrie dans la manière de présenter les trois variables. Z joue, dans les tableaux ci-dessus, un rôle très différent de celui de X et de Y. Dans le choix, on pourra se laisser guider par le nombre de valeurs de chaque variable, mais ce n'est pas suffisant. Le tableau ci-dessus se justifie particulièrement lorsque Y est la variable à expliquer, X la principale variable explicative, et Z une troisième variable susceptible de modifier l'influence de X sur Y.

Lorsque les deux variables explicatives jouent à peu près le même rôle, et que la variable expliquée (Y) est simplement dichotomique, et par exemple représente les réponses « oui » ou « non » à une question, on peut reporter dans chaque case du tableau XZ le pourcentage de « oui » de la façon suivante :

	Z	$\bar{Z}$
X	p_{11} %	p_{21} %
$\bar{X}$	p_{12} %	p_{22} %

P_{11} % représente le pourcentage de « oui » sur le sous-échantillon défini par les modalités X et Z, p_{21} dans le sous-échantillon $\bar{X}Z$, etc. Ce type de tableaux est très commode à lire, et on y repère facilement l'influence de chaque facteur, ainsi que la présence d'une interaction. Mais il faut garder en tête qu'il s'agit d'un tableau simplifié, et que Y n'y apparaît pas explicitement. S'il s'agit de présenter des résultats à un éventuel lecteur, cela devra être clairement indiqué, pour qu'il ne confonde pas ce tableau avec un simple tableau de contingence indiquant la relation entre X et Z.

Souvent, la simple inspection des pourcentages ne suffit pas : il est rare que tous ceux qu'on compare soient exactement égaux. Il est donc nécessaire de disposer d'instruments de mesure de la force d'une relation, et aussi de critères permettant de conclure qu'une relation apparente n'est pas significative. C'est la fonction des techniques statistiques.

La situation la plus commode, dans le cas de tableaux croisés, est celle où chaque variable n'a que deux modalités. On a alors des tableaux 2 x 2, faciles à analyser. On essaie donc souvent de se ramener à ce cas, en regroupant des catégories voisines. Lorsque les variables sont au moins ordinales, le seul problème est celui de savoir où placer la coupure entre les deux catégories finales. Lorsqu'il s'agit d'une variable nominale, comme il n'y a plus de proximités, on peut hésiter sur la composition même de chaque catégorie.

On peut aussi se ramener à une dichotomie en ne conservant que les catégories extrêmes, donc en négligeant tout ce qui se trouve entre elles. On a évidemment ainsi plus de chances de mettre en évidence une relation.

Ces diverses manières de se ramener à des tableaux 2 x 2 se justifient si l'on a des raisons d'admettre que la relation entre les deux variables est monotone. Sinon, il est évidemment nécessaire de conserver un nombre suffisant de valeurs pour mettre en évidence la forme de la relation.

Pour mesurer la *force* de la relation entre deux variables dichotomiques, on a proposé plusieurs coefficients. Le plus courant est le coefficient Q, dit de Yule.

Représentons les fréquences de chaque case par les lettres suivantes :

	X	$\overline{X}$
Y	a	b
$\overline{Y}$	c	d

la valeur de Q est donnée par :

$$Q = \frac{ad - bc}{ad + bc}$$

Dans le cas de relation symétrique parfaite, c'est-à-dire si une des diagonales est vide, $Q = 1$. S'il y a indépendance, $Q = 0$. En cas de relation inverse, si X est associé à $\overline{Y}$ et $\overline{X}$ à Y, $Q = 1$.

+ 1 et - 1 sont les deux valeurs extrêmes que peut prendre Q, et constituent donc les limites de son intervalle de variation. Plus la valeur de Q est élevée, plus on peut conclure que la relation est forte.

$Q = 1$ lorsque $b = c = 0$, donc lorsqu'il y a correspondance parfaite entre X et Y. Mais il suffit que l'une seulement de ces deux fréquences soit nulle pour que $b\ c = 0$, donc pour que $Q = 1$. Cette dernière valeur indique donc aussi bien l'existence d'une relation symétrique que d'une relation asymétrique parfaites. On ne peut, par conséquent, pas se contenter de connaître la valeur de Q, il faut encore regarder la répartition des valeurs dans le tableau pour savoir dans quel cas on se trouve.

La valeur de Q nous indique la force de la relation, donc la distance par rapport à l'indépendance des deux variables. Mais nous avons en général plutôt besoin de savoir s'il est préférable ou non de conclure à l'existence d'une relation dans la population parente. On utilisera pour cela des tests statistiques, le mieux adapté ici étant le test de χ^2, qui est suffisamment connu pour que nous ne l'exposions pas ici. Lorsque les fréquences sont faibles, on utilise de préférence le test de probabilité exacte de Fisher.

Dans l'interprétation des valeurs de Q et de χ^2, il faut prendre garde au fait qu'ils ne mesurent pas les mêmes grandeurs. Q mesure la force de la relation, χ^2 la distance par rapport à l'indépendance. La différence vient surtout du fait que la valeur de Q ne tient compte que des proportions, alors que le χ^2 fait intervenir la valeur des fréquences. On verra mieux ce point sur les exemples suivants :

	X	$\overline{X}$
Y	100	50
$\overline{Y}$	50	100

Q = 0,6
χ^2 = 33,2

	X	$\overline{X}$
Y	10	5
$\overline{Y}$	5	10

Q = 0,6
χ^2 = 3,32

La valeur de Q est la même dans les deux cas, alors que celle de χ^2 varie, et si on a décidé d'adopter le seuil de signification de 0,05, les conclusions qu'on en tirera seront différentes : on admettra que, dans le premier tableau il y a une relation, et pas dans le second.

Pourtant la « force » de la relation est la même. C'est que l'augmentation du nombre de sujets augmente la sensibilité du test. Des relations faibles (dans la population) seront plus probablement décelées avec un échantillon important qu'avec un échantillon faible. Ou, si l'on veut, de façon plus intuitive, on peut considérer que, dans le premier tableau, la relation qui se trouvait dans le second a été confirmée dix fois. On peut donc plus facilement conclure à sa réalité dans la population.

Lorsqu'on estime ne pas pouvoir regrouper les valeurs des variables pour se ramener à un tableau 2 x 2, l'analyse statistique est moins aisée, surtout lorsqu'une des variables est nominale, ou les deux. Il est théoriquement toujours possible d'utiliser le test de χ^2, mais il faut garder à l'esprit qu'on ne mesure ainsi qu'une distance par rapport à l'indépendance, dont les données peuvent s'écarter de diverses manières, qui pourront donner lieu à des interprétations très différentes. Il est donc, encore plus que dans le cas des tableaux 2 x 2, nécessaire de préciser la forme de la relation. Pour cela, il peut être utile de construire un tableau dans lequel on aura reporté dans chaque case la valeur de la différence observée (fo) et la fréquence attendue (fa), ou de préférence, les valeurs :

$$\frac{(fo - fa)^2}{fa}$$

dont la somme donne la grandeur de χ^2. En repérant les endroits, ou les zones, où ces valeurs sont fortes, on peut voir quelles sont les combinaisons de X et de Y qui sont particulièrement fréquentes ou rares, ce qui permettra de préciser la forme de la relation.

Si les deux variables sont quantitatives, on peut calculer le coefficient de corrélation de Bravais - Pearson. Il faut, dans ce cas, lors de l'interprétation, veiller à distinguer la valeur de ce coefficient, et la conclusion qu'on peut tirer concernant sa signification, et la valeur du coefficient de régression de Y sur X (ou de X sur Y).

Le coefficient de corrélation est la mesure de la *dispersion* du nuage de points autour d'une droite « moyenne ». Le coefficient de régression donne la valeur de la *pente de la droite* qu'on peut ajuster sur les moyennes des distributions partielles de Y à X constant. Un coefficient de corrélation élevé indique que, à chaque valeur de X ne corres-

pondent que des valeurs de Y peu dispersées. Un coefficient de régression élevé indique que Y varie fortement en fonction de X.

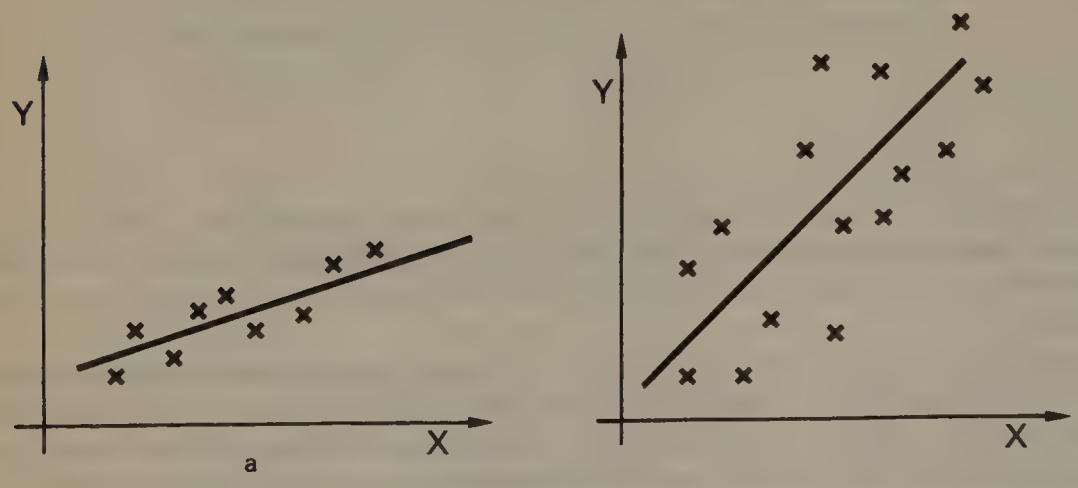


Fig. 15

Dans le cas a (fig. 15), le coefficient de corrélation sera élevé, les points représentant les observations étant peu dispersés autour de la droite de régression, mais la pente de celle-ci est faible, Y variant peu en fonction de X. La relation est faible, mais très nette.

En revanche, dans le cas b, c'est l'inverse, pour chaque valeur de X, la différence entre les valeurs de Y est élevée ; la relation n'est donc pas très stricte, ce qui permet de supposer l'intervention d'autres facteurs explicatifs de Y. Mais, en moyenne, Y varie beaucoup en fonction de X.

Cette distinction est surtout importante lorsqu'on analyse des relations entre trois variables, ce qui amène à comparer plusieurs tableaux croisant les mêmes variables. Lorsqu'on « contrôle » la relation entre X et Y par une troisième variable Z, c'est la comparaison des pentes des droites de régression de Y sur X dans chaque sous-échantillon défini par les modalités de Z qui est significative, et qui permettra de mettre en évidence l'existence d'une éventuelle interaction. Si l'on se contente de calculer les coefficients de corrélation, et de conclure, par exemple, qu'ils sont tous significatifs, on ne pourra pas déceler d'interaction que la comparaison des pentes pourrait mettre en évidence.

Lorsqu'on a calculé les coefficients de corrélation entre trois variables prises deux à deux, il est possible de calculer un *coefficient de corrélation partielle*, qui donne la valeur de la corrélation entre X et Y lorsque Z est maintenu constant. Toutefois, comme nous l'avons vu, rien n'impose que la relation entre X et Y soit la même dans toutes les modalités de Z. Le coefficient de corrélation peut être considéré comme une estimation

des valeurs de la corrélation entre X et Y dans chaque sous-échantillon. On ne peut admettre que ces valeurs sont égales que s'il n'y a pas interaction, ce qui est évidemment une hypothèse très solide. C'est donc une statistique à ne manipuler qu'avec beaucoup de prudence.

BIBLIOGRAPHIE

L'ouvrage d'ensemble le plus complet sur les problèmes de techniques de mesure dans les sciences humaines reste le livre déjà ancien :

- W.S. Torgerson, *Theory and Methods of Scaling*, New York, Wiley, 1958.

Depuis, les principaux progrès sont surtout apparus dans le domaine des techniques multidimensionnelles, dont il n'est presque pas question ici.

Citons encore, en français :

- R. Daval et al., *Traité de psychologie sociale*, vol. I, Paris, P.U.F., 1967.
- B. Matalon, *L'Analyse hiérarchique*, Paris, Mouton et Gauthier-Villars, 1965.
- P. Debaty, *La Mesure des attitudes*, Paris, P.U.F., 1963.

Pour des exemples d'utilisation d'indicateurs divers :

- R. Boudon, P. Lazarsfeld, *Le Vocabulaire des sciences sociales*, Paris-La Haye, Mouton, 1965.

Sur l'analyse des relations :

- R. Boudon, P. Lazarsfeld, *L'Analyse empirique de la causalité*, Paris-La Haye, Mouton, 1966.
- M. Rosenberg, *The Logic of Survey Analysis*, New York, Basic Books, 1968.
- J.A. Davis, *Elementary Survey Analysis*, Englewood Cliffs, Prentice Hall, 1971.
- R. Boudon, *Les Mathématiques en sociologie*, Paris, P.U.F., 1971.

Sur l'analyse algébrique :

- C. Flament, *L'Analyse booléenne de questionnaire*, Paris-La Haye, Mouton, 1976.

Sur l'analyse de données :

- J.-P. Benzecri, *L'Analyse de données*, Paris, Dunod, 1973.

Sur les méthodes graphiques :

- J. Bertin, *Sémiologie graphique*, Paris-La Haye, Mouton, 1972.

TABLE DES MATIÈRES

1. PROBLÈMES GÉNÉRAUX 5

1.1 *Définitions et généralités* 5
1.2 *Enquête, observation, expérimentation* 10
1.2.1 Définitions 11
1.2.2 Relations entre les méthodes 11
1.2.3 Utilisation des enquêtes 15
1.3 *Conclusion* 19

PREMIÈRE PARTIE : LA PRODUCTION DES DONNÉES D'ENQUÊTES

2. QUI INTERROGER ? 25

2.1 *Le choix de la population à interroger* 25
2.2 *Les méthodes d'échantillonnage* 29
2.2.1 Les échantillons représentatifs 29
2.2.1.1 Définition, 29. - 2.2.1.2 La constitution des échantillons aléatoires, 30. - 2.2.1.3 Unités et grappes, 33. - 2.2.1.4. Les échantillons stratifiés, 35.
2.2.2 Les échantillonnages sur place 36
2.2.3 Les échantillonnages par quota 38
2.2.4 Les biais dans la réalisation de l'échantillon 41
2.2.4.1 Vue d'ensemble, 41. - 2.2.4.2 Les biais d'échantillonnage dans les échantillons aléatoires, 42. - 2.2.4.3 Les biais d'échantillonnage dans les enquêtes par quota, 46. - 2.2.4.4 L'importance des biais de réalisation de l'échantillon, 47.
2.2.5 L'échantillonnage pour les enquêtes non statistiques 50
2.2.6 La taille de l'échantillon 51
2.2.7 Est-il vraiment nécessaire qu'un échantillon soit représentatif ? 53

3. COMMENT INTERROGER ? LES ENTRETIENS 57

3.1 *Les différents modes d'interrogation* 57
3.2 *Caractéristiques générales de l'entretien* 59
3.3 *La situation d'entretien : Concepts mis en jeu* 61

3.3.1 Facteurs liés à la situation 63
3.3.2 Facteurs liés à l'interviewé 64
3.3.3 Facteurs liés à l'enquêteur 68
3.3.4 Facteurs liés au langage 71

3.4 *Les différentes sortes d'entretiens* 73

3.4.1 Différents types d'entretiens 75
3.4.2 Quel entretien pour quel problème ? 76

3.5 *Les techniques utilisées lors des entretiens* 80

3.5.1 Etablissement de la relation E.-e. 80
3.5.2 La formulation de la consigne 81
3.5.3 La relation E.-e. en cours d'entretien 85

3.6 *Le choix, la formation et le contrôle des enquêteurs* 89

3.6.1 Le choix des enquêteurs 89
3.6.2 La formation des enquêteurs 91
3.6.3 Le contrôle des enquêteurs 91

4. COMMENT INTERROGER ? LES QUESTIONNAIRES 93

4.1 *Approche qualitative et approche quantitative* 93
4.2 *Les grandes lignes d'un questionnaire* 98
4.3 *Les différents types de questions* 101
4.4 *Questions ouvertes et questions fermées* 102
4.5 *La formulation des questions* 106

4.5.1 Problèmes généraux 106
4.5.2. Les questions de fait 110

4.5.2.1 Questions de fait et questions de connaissance, 110. - 4.5.2.2 Les événements uniques, 111. - 4.5.2.3 Les événements répétitifs, 111.

4.5.3. Les questions d'opinion 116

4.5.3.1 Problèmes généraux, 116. - 4.5.3.2 La forme des questions et des réponses, 118. - 4.5.3.3 Le libellé des questions, 122. - 4.5.3.4. L'organisation des questions, 127. - 4.5.3.5 Quelques thèmes particuliers, 129.

4.6 *Les pré-tests du questionnaire* 136

5. LE DÉROULEMENT D'UNE ENQUÊTE 139

5.1 *Les conditions de passation* 139

5.1.1 Les lieux 139
5.1.2 Le contact 140
5.1.3 Le déroulement 142
5.1.4 Les questionnaires auto-administrés et les enquêtes par correspondance 143

5.2 *Esquisse d'une psychosociologie de la situation d'enquête* 145

5.2.1 La situation d'enquête comme situation sociale 145
5.2.2 Les rôles de l'enquête 148
5.2.3 Les objectifs de l'enquêté 149
5.2.4 L'influence de l'enquêteur 150

DEUXIÈME PARTIE : L'ANALYSE ET L'INTERPRÉTATION DES DONNÉES D'ENQUÊTE

6. L'ANALYSE DE CONTENU 155

Généralités 155

6.2 *Procédures closes et ouvertes* 158

6.3 *Champ d'application de l'analyse de contenu dans une situation d'enquête* 160

6.4 *Méthodologie de l'analyse de contenu* 163

6.4.1 Comment coder ? 164

6.4.1.1 Catégories d'analyse, 164. - 6.4.1.2 Unités d'enregistrement, de contexte, de numération, 167.

6.4.2 Fiabilité de l'instrument ? 170

6.4.2.1. Fidélité, 170. - 6.4.2.2 Validité, 171.

6.5 *Quelques méthodes d'analyse de contenu* 172

6.5.1 Procédure close 173

6.5.1.1 Catégories psychologiques, 173. - 6.5.1.2 Catégories psycholinguistiques, 175. - 6.5.1.3 Catégories psychosociologiques, 178. - 6.5.1.4 Catégories linguistiques, 183.

6.5.2 Procédures ouvertes 184

6.5.2.1 Le décompte fréquentiel, 184. - 6.5.2.2 L'analyse thématique, 185. - 6.5.2.3 Les concomitances thématiques, 186. - 6.5.2.4 L'analyse par grappes, 186. - 6.5.2.5 L'analyse par champs sémantiques, 192.

6.6 *L'agrégation des analyses* 193

6.7 *Conclusion* 202

7. L'ANALYSE STATISTIQUE DES QUESTIONNAIRES 205

7.1 *Les traitements préliminaires* 205

7.1.1 La mise en forme des résultats 205

7.1.2 La codification des questions ouvertes 209

7.1.3 Les questions à réponses multiples 211

7.1.4 Les non-réponses 216

7.2 *Les variables* 219

7.2.1 La construction des variables 219

7.2.1.1 Observables, indicateurs et variables intermédiaires, 219. - 7.2.1.2 La construction des échelles d'attitude, 225.

7.2.2 Les variables dites classiques 246

7.2.2.1 Intérêt et limites de l'emploi de ces variables, 246. - 7.2.2.2 Les principales variables classiques, 251.

7.2.3 Les tris à plat et l'estimation des grandeurs 260

7.3. *Les relations* 262

7.3.1 Définitions 262

7.3.2 L'analyse des relations binaires 263

7.3.2.1 Le langage des relations, 263. - 7.3.2.2 Les relations entre observables, 264. - 7.3.2.3 L'interprétation des relations, 266.

7.3.3 Les relations entre trois variables 269

7.3.3.1 La description des relations, 269. - 7.3.3.2 L'interprétation des relations entre trois variables, 274.

7.3.4 Les relations entre plus de trois variables 279

7.3.4.1 L'analyse de causalité, 279. - 7.3.4.2. La méthode de segmentation, 281. - 7.3.4.3 L'analyse des structures de réponses, 282. - 7.3.4.4 Remarques sur l'analyse des données, 283.

7.3.5 La lecture et la présentation des tableaux croisés 288

Coventry University

Armand Colin Éditeur
5, rue Laromiguière, 75241 Paris Cedex 05
N° 105700
Dépôt légal : juillet 1995

SNEL S.A.
Rue Saint-Vincent 12 - 4020 Liège
juillet 1995

Qu'il s'agisse d'enquêtes de marché ou de sondages d'opinion, l'enquête est l'instrument clé de toutes les investigations psychosociologiques.

Ce livre propose aux praticiens et aux étudiants l'acquisition des « savoir-faire » indispensables. Il suit, dans son déroulement, les différents moments de l'enquête : choix d'une population, échantillonnage, entretiens et questionnaires, analyse de contenu… Pour chacune de ces étapes, sont indiqués les méthodes existantes, les outils disponibles, et précisés leur fonctionnement aussi bien que les limites de leur utilisation.

Mais la diversité des types d'enquêtes comme des techniques auxquelles elles ont recours implique une réflexion théorique sur chacune de ces opérations, qui permette le choix d'une méthode appropriée. Cette approche méthodologique, liée constamment à l'analyse de la pratique, est sans doute l'un des éléments les plus originaux de cet ouvrage.

Une enquête consiste à susciter un ensemble de discours individuels, à les interpréter et à les généraliser, c'est-à-dire à les réduire à un discours unique : un rapport de fin d'étude, par exemple, qui livre des résultats. Se trouve ainsi posé, à chaque instant, un problème de « sens ».

Rodolphe Ghiglione est professeur de psychologie à l'université de Paris VIII où il est responsable du groupe de recherches sur la parole. Il est l'auteur, avec divers collaborateurs, du Manuel d'analyse de contenu, *paru dans la collection U.*

Benjamin Matalon, qui a enseigné à l'université de Genève, est professeur de psychologie à l'université de Paris VIII. Il a collaboré aux travaux d'organismes et de sociétés dont les activités « sur le terrain » sont bien connues (SEMA, CREDOC, etc.) et il est l'auteur d'un ouvrage sur L'analyse hiérarchique.

KN-823-135

ISBN: 2-200-21635-1

9 782200 216351

Maquette et illustration Guy Sabadie.